

Department of Public Health, University of Helsinki, Finland

ECONOMIC EVALUATION OF SELECTED FINNISH SCREENING PROGRAMMES

Tiina Leivo

Academic Dissertation

To be publicly discussed, by permission of the Medical Faculty of the University of Helsinki, in the Auditorium XII in the main building of the University of Helsinki, on August 3rd, 2001, at 12 o'clock noon.

Helsinki 2001

Supervised by :

Professor Harri Sintonen

Department of Public Health

University of Helsinki

and

Docent Risto Tuominen

Department of Public Health

University of Helsinki

Reviewed by:

Associate Professor Ivar Sonbo Kristiansen

Institute of Public Health, Health Economics Research Unit

University of Odense, Denmark

and

Docent Pekka Rissanen

Department of Public Management

University of Vaasa

Opponent:

Docent Jouko Isolauri

Department of Surgery

Turku University Central Hospital

ISBN 952-91-3614-5

ISBN 952-10-0066-X (pdf version, available at <http://ethesis.helsinki.fi>)

Yliopistopaino

Helsinki 2001

To Janne, Joel, Julia and Jasmin

CONTENTS

LIST OF ORIGINAL PUBLICATIONS	7
I INTRODUCTION	9
II ECONOMIC EVALUATION	13
III REVIEW OF THE LITERATURE	20
3.1 Economic evaluation of screening programmes in Finland	20
3.2 Economic evaluation of ultrasound screening in pregnancy	20
3.3 Economic evaluation of breast cancer screening	21
3.4 Economic evaluation of double - reading mammograms	23
3.5 Economic evaluation of H pylori screening	24
IV AIMS OF THE STUDY	26
V MATERIAL AND METHODS	27
5.1 Effectiveness data (I, II, III)	27
5.2 Cost data (I, II, III)	30
5.3 Measure of cost-effectiveness (I, II, III)	32
5.4 Costs and benefits in cost-benefit analysis (IV)	33
5.5 Valuation and discounting (I, II, III, IV)	39
5.6 Sensitivity analysis (I, II, III, IV)	40
VI RESULTS	43
6.1 Ultrasound screening in pregnancy (I)	43
6.2 Breast carcinoma screening (II)	44
6.3 Double-reading mammograms (III)	46
6.4 Helicobacter pylori screening (IV)	47

VII DISCUSSION	49
7.1 Material and methods (I, II, III, IV)	49
7.1.1 Cost-effectiveness-analysis material and methods (I, II, III)	49
7.1.2 Cost-benefit-analysis material and methods (IV)	51
7.1.3 Valuation and discounting methods (I, II, III, IV)	53
7.2 Ultrasound screening in pregnancy (I)	54
7.3 Breast carcinoma screening (II)	55
7.4 Double-reading mammograms (III)	57
7.5 Helicobacter pylori screening (IV)	58
 VIII SUMMARY AND CONCLUSIONS	 59
 APPENDIX 1 CURRENT SCREENING PROGRAMMES IN FINLAND	 62
APPENDIX 2 PROPOSED SCREENING PROGRAMMES IN FINLAND	65
 IX ACKNOWLEDGEMENTS	 67
 X REFERENCES	 69

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following articles, which are referred to in the text by the Roman numerals I-IV.

I Leivo T, Tuominen R, Saari-Kemppainen A, Ylöstalo P, Karjalainen O, Heinonen O.P. Cost-effectiveness of one-stage ultrasound screening in pregnancy: a report from the Helsinki ultrasound trial. *Ultrasound Obstet. Gynecol.* 1996;7:309-14.

II Leivo T, Sintonen H, Tuominen R, Hakama M, Pukkala E, Heinonen O.P. The Cost-effectiveness of Nationwide Breast Carcinoma Screening, 1987-1992 in Finland, 1987-1992. *Cancer* 1999;86:638-46.

III Leivo T, Salminen T, Sintonen H, Tuominen R, Auerma K, Partanen K, Saari U, Hakama M, Heinonen O.P. Incremental Cost-Effectiveness of Double-reading Mammograms. *Breast Cancer Res. & Treatment.* 1999;54:261-7.

IV Leivo T, Salomaa A, Kosunen T, Tuominen R, Färkkilä M, Linna M, Sintonen H. Cost – benefit analysis of *Helicobacter pylori* screening. Submitted.

I INTRODUCTION

In order to optimize the use of scarce resources, economic evaluation of health interventions is a necessity highly desirable. From the viewpoint of national economy, optimization leads to more utility per resource unit used.

Economic evaluation is an aid to decision making, but criteria other than efficiency may also be relevant when making value judgements as to whether or not to provide a particular screening programme. In particular, concerns of equity, which are not formally addressed in most economic analyses, may have an important impact on decisions (Brown et al 1998).

Other methods than economic evaluation are in practice often used when choices about health interventions are made. Methods based on notions of medical necessity or standards of evidence are used. However, these methods often implicitly involve both economical and value judgements, but in fact lack the information about costs and health consequences (Gold et al 1996).

The aim of screening is to detect a disease in preclinical phase to reduce future mortality and morbidity. From a broad perspective a considerable part of health care can be defined as preventive. Especially if psychiatric conditions are also taken into account, almost every doctor-patient visit could be defined as preventive. Preventive interventions fulfil the requirements of a screening programme, if a specified screening programme protocol, including a defined screening test exists. A screening test can also be a formulated question sheet to screen for behavioural or psychiatric diseases.

An ideal screening programme should have the following medical prerequisites (Wilson et al 1968, Hakama et al 1990, Chamberlain 1984, Day et al 1984, Hristova 1996):

A disease is suitable for screening if:

- it is an important health problem justifying the efforts of screening
- there is a long enough detectable asymptomatic preclinical phase
- the proportion of lesions found in this preclinical phase that would progress to clinical lesions is significant
- an acceptable treatment is available which can improve a patient's prognosis for the earlier diagnosis

Modifications of these criteria encompass the need of economic evaluation of screening programmes in that they recognise the economic costs of the programme

have to be considered in relation to the benefits of early detection. Generally accepted screening criteria include also (Frame 1986):

- tests that are acceptable to patients must be available at reasonable cost to detect the condition in the asymptomatic period
- the incidence of the condition must be sufficient to justify the cost of screening

The test should be valid and identify the disease in its preclinical phase. The test should be acceptable for the population - easy to apply, painless and without side effects.

A screening programme should have the following features:

- the target population should be identified
- individuals in the target population should be identifiable
- high coverage and attendance rates can be guaranteed
- there should be facilities for conducting the tests
- there should be a designed and agreed referral system
- there should be an organized quality control system

At present, mostly nationwide universal screening programmes are run in Finland. In a nationwide programme the target population is identified by age, gender or pregnancy. Another option would be to identify the target population by relative risk of morbidity, this is less common.

In Finland preventive medicine has been emphasized. As a consequence, relatively many nationwide screening programmes are being undertaken. All these programmes are free of charge for the attendants according to the principle of equity. The attendance rates in all Finnish screening programmes have traditionally been very high. At present Finnish nationwide screening programmes include:

1. Antenatal screening

Finland has a very intensive preventive maternity welfare programme which is organised by the local health centres' maternity counselling clinics. According to the national guidelines primiparas should visit the maternity guidance center 13-17 times, and multiparas 9-13 times respectively during pregnancy. The antenatal screening programme includes several physical, clinical and laboratory screening tests to detect pregnancy-induced diseases or disorders, infections, developmental abnormalities and malformations. Appendix 1. (Hiilesmaa et al 2000, Viisainen K 1999, Leinikki et al 2000):

2. Perinatal screening

Perinatal screening is performed in the hospital where the child is born including laboratory and clinical examinations. Appendix 1.

3. Child counselling clinic screening

Under seven-years-old children are offered several child counselling clinic visits (Simell et al 1997, Yleislääkärin käsikirja 1994). The child guidance clinic programme includes 6 GP visits and around 14-18 nurse visits. The visits include guidance, a vaccination programme and numerous specified clinical screening tests. Appendix 1. Screening for behavioural or psychological disorders is also recommended, but no formal screening tests, i.e. question sheets, are used.

4 Screening in the defensive armed forces

Finland has a liability to military service. Four physical doctor's examinations are due before and during a term of the service. The aim of the examinations is to define the military service class and to prevent morbidity during the service. Appendix 1.(Lääkärintarkastusohje 1997, Yleislääkärin käsikirja 1994):

The specified question sheet includes also questions about previously diagnosed diseases, which do not fulfil the definition for screening.

5 Screening in occupational health care

Screening in occupational health care is common, although no formal screening programmes exist. Most often the routine testing of blood pressure, serum lipids and cholesterol, hemoglobin, blood glucose and small blood count take place.

6 Cancer screening

Finland has two nationwide cancer screening programmes. All 30-60-years-old women are invited every fifth year to cervical cancer screening. All 50-59 years-old women are invited every second year to breast cancer screening (Teppo et al 2000).

At present, targeted non-nationwide screening is organised for chronically ill patients. For example, diabetes mellitus- patients are offered several screening tests during the course of the disease. Adenoma – and ulceritis colitis patients are offered targeted colorectal cancer colonoscopic screening.

There is continuous pressure from health care professionals and from the public to introduce new screening programmes to prevent morbidity and mortality. The health care professionals see the possibilities of new technologies. For example, at present several hundred genetic diseases could be screened by gene technology (Suomen Lääkärilehti, Uutispäivävyri 1997). At the level of national economy, it has been

observed that the higher an individual's income, the proportionally higher is their expenditure for health care.

Currently in Finland, several new screening programmes are under discussion or in their pilot phase (Hiilesmaa et al 2000, Leinikki et al 2000, Teppo et al 2000, Saarelma et al 2000, Robinson 1998). Appendix 2.

In this study two nationwide Finnish screening programmes, and one pilot Finnish screening programme aimed at nationwide use are subjected to economic evaluation.

In Finland, 65 000 women become pregnant annually. Ultrasound screening in pregnancy is used in most Finnish communities aiming at early detection of malformation and the subsequent termination of the pregnancy, if the mother so wishes.

Breast carcinoma is the most common cancer affecting the female population in Finland. The mean annual number of new cancer cases diagnosed in 1988-1992 was 2647. (Finnish Cancer Registry 1995) Breast carcinoma screening, covering women ages 50-59 years is used nationwide in Finland aiming at the detection of cancer at an early and curable stage. The double reading practice of mammograms is the strategy actually used in mammography screening in Finland. The double reading practice is recommended aiming at better detection rate of cancers within about half of the screening programmes internationally (Ballard-Barbash 1999).

Helicobacter pylori infection is typically acquired in childhood. It causes bacterial infection of the gastric mucosal surface that causes progressive damage of the stomach. The possible clinical manifestations such as peptic ulcer and gastric malignancy appear in adulthood. (Axon et al 1997, Blaser 1990). The prevalence of the *H. pylori* infection is a cohort- phenomenon, being 80-90% in the cohorts born in the beginning of the century and 10-20% in the cohorts born in the 1970's (Sipponen 1995). The infection can be treated with antibiotics. *Helicobacter pylori* pilot screening is used in one Finnish community. The population – based Vammala *H. pylori* pilot screening study was started in 1994 in Vammala, a semi-urban community in South-West Finland. Until the end of 1998, 5288 screenees in the 15 to 45 age group have entered the programme (Salomaa et al 1998).

II ECONOMIC EVALUATION

Economic evaluation can be defined as the comparative analysis of alternative courses of action, e.g., to screen or not to screen, in terms of both their costs and consequences. The use of any scarce resources, be they manpower, buildings or equipment, has an opportunity cost in terms of the benefits foregone by denying those resources to other competing claims. Economics is concerned principally with allocating resources efficiently. Efficiency is about making choices, which derive the maximum total benefit from the limited resources available (Drummond 1987, Brown et al 1998).

The principal methods for full economic evaluation are: cost-minimization analysis, cost-effectiveness analysis, cost-benefit analysis and cost-utility analysis. The identification of various types of costs and their subsequent measurement in monetary terms is similar across these different evaluation methods. However, the nature of the consequences stemming from the alternatives being examined may differ considerably. (Drummond 1987).

An economic evaluation through cost-minimization analysis requires that no outcome differences exist between the alternatives. In most circumstances, however, the outcomes will differ in some way and then the technique that can consider differences in both costs and outcomes is required (Drummond 1987, Brown et al 1998). Cost-minimization analysis is really a special form of cost-effectiveness analysis, where the consequences of the alternative treatments being compared turn out to be equivalent (Drummond et al 1997).

Cost-effectiveness analysis requires a single effect of interest, common to both alternatives, but achieved to different degrees. The measurement of consequences is in natural units, e.g., life years gained. In assessing screening programmes, process measures such as the proportion of cases detected are often used as the measure of effectiveness. The additional costs and effects of an alternative programme are often presented in terms of an incremental cost-effectiveness ratio. In cost-effectiveness analysis comparison cannot be made across health care programmes, where different outcome measures are used. Moreover, it is unlikely that all the important outcomes are captured by a unidimensional measure of effectiveness (Drummond 1987, Brown et al 1998).

In cost-benefit analysis effects can be single or multiple, not necessarily common to both alternatives, and common effects may be achieved to different degrees by the alternatives. The measurement of consequences is uniquely in monetary terms. In practice complete cost-benefit studies are rarely undertaken to evaluate health care programmes, because of the practical difficulty and the social dislike of putting monetary values on life and suffering (Drummond 1987, Brown et al 1998). The

recent literature contains a number of studies that assess individuals' willingness to pay for health benefits. A comprehensive cost-benefit analysis of health care interventions would use this approach to value health benefits. However, very few cost-benefit analyses incorporating these estimates have so far been published. (Drummond et al 1997).

In cost-utility analysis the effects are measured in quality-adjusted life-years gained. It provides a broader comparability between different programmes than cost-effectiveness analysis. Programmes can be ranked according to their incremental (additional) cost per QALY gained and, in the context of a fixed budget for health care, those programs offering additional QALYs at lowest additional cost per QALY should be given priority (Drummond 1987, Brown et al 1998).

Economic evaluation is according to Brown et al (1998) particularly relevant to screening for a number of reasons. Firstly, screening is discretionary: the case for intervention can be relatively dispassionately decided in the light of the probabilistic expectation of benefit to a proportion of those screened. Screening is not normally subject to the pressures from the social 'rule of rescue'. Secondly, within any screening programme a number of different strategies exist. Thirdly, formal screening programmes require substantial investment in infrastructure. Economic evaluation can contribute to decisions whether a new screening programme for a particular disease or disorder should be used, and can also aid in decisions about changes to existing programmes.

Screening in health care can be viewed as an investment, which yields additional information. According to Cairns et al (1993) screening has a number of distinctive features: the range of potentially relevant outcomes is broader, the proximate investment is in knowledge and not health, patients undergoing screening are generally asymptomatic.

Economic evaluation of screening has also a number of distinctive key issues in the measurement and valuation of costs and outcomes, sources of data, discounting and sensitivity analysis. The economic evaluation of antenatal screening raises additional problematic issues.

Economic evaluation studies adopt a broad societal perspective, i.e. the analyst considers everyone affected by the intervention and counts all significant health costs and outcomes that flow from it, regardless of who experiences the outcomes or costs. Implicit in the societal perspective is the recognition that societal resources are limited and that health should not be exempted from these limits. (Gold et al 1996). Broad perspective means also that the screening is not regarded solely as the test itself defining positive or negative result, but as a diagnostic and health care process beginning from the screening test (Sintonen 2000).

According to Gold et al (1996) health intervention, e.g. screening, imposes changes in health care, non-health care and time resources. Health care costs include the downstream or negative costs (or savings) associated with the intervention. Drummond et al (1997) divides the resource use costs by sectors 1 health care 2 patient and family 3 other sectors (e.g. other public agencies). He classifies the resources saved as part of the consequences, but reports that they mirror the costs and are measured and valued in a similar way. In the previous literature costs and benefits were often classified as direct, indirect and intangible. Terms direct costs and benefits are used to denote the resources consumed (costs) or saved (benefits) by the programme. In the main these would be resources in the health care sector. The terms indirect costs and benefits are used to denote the time of patients (or their families) consumed or freed by programme. Terms intangible costs and benefits are used to denote those consequences that are difficult to measure and value (Drummond et al 1997).

Screening costs are typically partly imposed on the individuals concerned. Costs incurred by individuals are likely to have an important effects on behaviour, particularly attendance rates (Brown et al 1998). The cost to society of the time taken to use health care may differ from the private opportunity cost of such time. Some researchers have emphasized that the private opportunity cost of time is a better predictor of demand for a screening service than the societal approach (Torgenson et al 1994). The valuation principles of the opportunity cost of unpaid inputs are under debate among health economists (Posnett et al 1996). Productivity costs are the costs associated with a loss or impaired ability to work or engage in leisure activities due to morbidity or death. The inclusion of productivity costs is also under debate among health economists. According to Gold et al (1996) productivity costs should not be included in the cost accounting, as they are included in the denominator of a cost-utility ratio, e.g. as quality-adjusted life-years saved. This has been challenged by other research groups (Brouwer et al 1997).

The inclusion of health and non-health care costs for unrelated illness in added years of life is an unsolved theoretical issue in health economics. Due to the long time-horizon it is of importance in the economic evaluation of screening. For example, would decision makers wish to oppose a smoking prevention program on the basis of a cost-effectiveness analysis that included the costs of future health care for individuals who were spared from a premature death of lung cancer and other smoking-related illness? (Gold et al 1996). For example, Pekurinen (1992) has reported that the estimated social costs slightly exceed the social benefits of smoking, and that it seems evident, that smoking is profitable to the public sector, apart from local authorities. In addition there are practical difficulties in including costs for unrelated illness in added years of life. Existing data may not be adequate to capture future resource use of all unrelated diseases and it may be unduly difficult to ascertain the effect of an intervention on the range of future causes of morbidity and death (Gold et al 1996).

Typically, economic studies evaluating screening have focused on the outcomes associated with true positive findings. The major change in outcomes associated with screening will be experienced by this group. For most screening situations, true positive results will bring forward the time of detection and allow earlier treatment (Brown et al 1998). However, screening unavoidably also medicalizes the screenees (Rose 1992). In screening, the test actually places the people into four categories; true positive, false positive, true negative and false negative according to whether or not they have the disease and whether or not the test result is positive or negative (Thorner et al 1961). The monetary values of the different possible results of the different test results should also be estimated (Simpson et al 1978). There may also be a small risk associated with screening and some pain or discomfort directly associated with the test or screen.

A false positive screening result will result in anxiety. Anxiety may be short-lived if the results are negative, but may remain for several months, or in some cases years, after a false positive result. Anxiety may also be experienced by non-attenders; family and friends. All these impacts of false positive screening results have been reported e.g. in mammography screening programmes. Negative screening result is usually reassuring. False negative screening results may not simply fail to bring forward detection, but may provide false reassurance, thereby delaying subsequent clinical diagnosis. (Brown et al 1998, Marteau 1990, Stewart-Brown 1997, MacLean et al 1984, Gram et al 1993).

In screening, information per se may also be valued for its own sake (non-decisional value) in addition to its decisional value. Berwick et Weinstein (1985) investigated the role of information in a study of willingness-to-pay for ultrasound in normal pregnancy. The average respondent allocated 74 percent of her willingness-to-pay to the decisional aspect of the information and 26 percent to information which was not used for decision making. Lange et al (1990) have reported similar results in a cost-benefit analysis of prenatal screening for autosomal dominant polycystic kidney disease.

When assessing the efficacy of an intervention the randomised, double blind controlled trial has the highest internal validity, that is, freedom from bias. In practice, the blinding is often very difficult to organise in clinical trials. However, as long as the data are drawn from clinical trials, it always lacks knowledge of the real life situation, i.e., the clinical trials tend to reach better efficacy than nationwide interventions mainly due to more strict clinical protocols and more resources in the organisation of the intervention. Effectiveness data from meta-analyses might have the advantage that the confidence interval around the point estimate of clinical effect is usually narrower than that from an individual trial and the result may be more generalisable (Drummond et al 1996). However, in practice possibilities to set up large and long-term randomised controlled trials are infrequent, not least because of the costs entailed. Where a clinical trial is not feasible, available data can be synthesised by modeling techniques. In such circumstances the economic evaluation is subject to some uncertainty. Even when economic evaluation has been incorporated in

a clinical trial, some modeling is likely to be necessary to allow for differences between the trial participants and the target population, to extrapolate from short-term outcomes to long-term survival and quality of life. Modeling studies undertaken before a clinical trial is initiated can help to identify the key parameters that need to be estimated within a planned trial (Brown et al 1998).

Economic analysis takes into account the timing of costs and outcomes. Individuals and society prefer resources now rather than later and would prefer to postpone costs. This is allowed for in economic evaluation by discounting future costs and outcomes to estimate their present value to us now. The rate of discount will vary between societies and over time. Consequently it varies between different studies. An important issue is, whether the outcomes should be discounted at the same rate as costs. Most researchers use the equal rate for costs and outcomes. The most often recommended rate is 3% or 5% per annum. However, for example, in Britain the Department of Health now recommends that life years and QALYs should be discounted at 1.5-2%, and costs at 6 %, respectively. The discount rate for costs and outcomes is of great significance for screening, where invariably the costs will occur before the outcomes, and where sometimes the outcomes may only arise after several decades. The higher the rate of discount applied to outcomes the less attractive screening will appear, particularly when compared with those treatment interventions that give immediate outcomes (Gold et al 1996, Department of Health 1995, Brown et al 1998, van Hout 1998).

In economic analyses uncertainty about estimates of costs, effectiveness, and the C/E ratio can arise in a number of ways: Two major sources of uncertainty are 1) parameter uncertainty and 2) model uncertainty. Parameter uncertainty is uncertainty about the true numerical values of the parameters used as inputs. Model uncertainty includes both uncertainty about the correct method for combining these parameters and uncertainty introduced by the combination of decisions made by an individual analyst (Gold et al 1996).

The standard way of dealing with parameter uncertainty is to use sensitivity analyses. Sensitivity analyses should be an essential part of a thorough economic evaluation, although at present all studies do not give an adequate account of uncertainty (Briggs et al 1995). Sensitivity analyses are mathematical calculations that isolate factors involved in an economic analysis to indicate the degree of influence each factor has on the outcome of the entire analysis (Gold et al 1996). It measures the uncertainties of the analysis. Univariate sensitivity analysis examines one variable, multivariate analysis multiple variables, respectively, at a time. In screening, the sensitivity analyses are especially important due to the long time horizon and extensive use of modeling and discounting.

Also statistical methods can be used, if the source of the uncertainty is the sampling variation in estimates of the parameters used in the analysis. Currently, there are three ways to provide an estimate of the uncertainty in the estimate of the cost-effectiveness

ratio. The first relies on the delta method to calculate the variability of any composite measure. The second relies on simulating the variance of the estimated C/E ratio or the distribution of the estimated C/E ratio, based on estimates of the variance-covariance matrix of the parameter estimates. The third is to derive a bootstrap estimate of the probability distribution of the ratio, its confidence interval, or the variance in the ratio (Gold et al 1996).

Model structure uncertainty is difficult to incorporate formally. If the magnitude of the error that might be introduced by structural assumptions is significant, the qualitative remedy appears to be sensitivity analysis, computing the C/E ratio estimate under the alternative assumptions and reporting the magnitude of the effect. For quantitative analysis, a weighed average of the resulting estimates might be computed, with weights reflecting the degree of confidence in each structural form (Gold et al 1996).

The economic evaluation of antenatal screening is generally more complex than evaluation of programmes, which screen individuals for conditions which might only affect those individuals. Antenatal screening raises the issue of saved resources, or averted costs, associated with the termination of affected pregnancies. Had the affected pregnancy not been detected through screening and terminated, costs would have been incurred throughout the affected child's lifetime to treat the condition for which it was screened. These are costs over and above the cost of a 'replacement' child without the condition. If a health service perspective is adopted, the excess costs avoided can be considered equivalent to the health service costs of treating the condition which would otherwise have been detected clinically at a later stage.

When the resource use implications for other sectors of society are considered the issue comes more complicated: for example, the avoided excess costs associated with educational and institutional care, would need to be considered, as well as the costs of voluntary services and care incurred by the family. (Cairns et al 1993, Brown et al 1998).

Some researchers have also tried to quantify the intangible psychological costs to the parents from having a handicapped child compared with 'the greater joy' they might experience from having a 'replacement' non-handicapped child, the opportunity for the latter arising from the existence of the screening programme (Hendersson 1982).

Antenatal screening also raises many ethical questions. In Finland, in 1985, the lutherian state church appointed a multi-disciplinary expert panel on ethical issues concerning antenatal screening, infertility treatments and genetics. In their report, selective abortion was the main ethical issue of antenatal screening. Selective abortion can be defined as an induced abortion of a handicapped fetus. From the christian perspective the values of life should not uniquely be effectiveness, success, ease and vitality, but the limited capabilities of a human being also have an important meaning. All medical diagnostics and treatments should be based on the rights of voluntariness and self-determination. According to the panel, in principle three kinds of arguments

can be stated for the selective abortion: 1) The life of the child would include unreasonable suffering. 2) The life of the parents would include unreasonable suffering. 3) The economical and population-political reasons justify selective abortions. In their report, the panel concluded, that only the suffering of parents should justify selective abortion (Kirkon yhteiskunnallisen työn keskus 1987).

Economic evaluations of screening programmes are population- and health care system specific. The disease and participation rates are variable among different populations, having a strong effect on the outcome of an economic evaluation. The participation rate is often dependent on the health care system thorough financial and organizational features, i.e., the possible fee of the screening, methods of invitation and further communication. The quality of the screening cascade differs also in different health care systems. Some researchers have proposed methods for carrying out country-specific economic evaluation from multinational or second country-specific data (Willke et al 1998, van Ineveld et al 1993).

Decision indices are developed for decision making to evaluate choices about health interventions. In cost-effectiveness-analysis, cost-effectiveness ratios are computed from the societal perspective and placed in rank order. A decision maker can select the intervention with the lowest cost per quality-adjusted life-year and continue down the list selecting interventions, until the available funds are exhausted. In cost-benefit analysis benefit- cost ratios are calculated, and an alternative with a higher benefit-cost ratio is preferred. The second common option is to select an alternative with a higher net benefit or net present value. However, decisions in the real world are more complicated. Decision indices provide valuable information about tradeoffs in the broad allocation of health resources, but other factors need to be considered as well – concepts of fairness and justice and practical questions of feasibility, that are not fully captured in the decision indices (Gold et al 1996, Sintonen et al 1997).

III REVIEW OF THE LITERATURE

3.1 Economic evaluation of screening programmes in Finland

Economic evaluation of the Finnish screening programmes is scarce. Of the ongoing Finnish screening programmes breast cancer (Hristova et al 1996) and cervical cancer screening have been preliminarily evaluated (Hristova et al 1996). No evaluation has been conducted for ultrasound screening during pregnancy in Finland.

Of the potential Finnish screening programmes first-void urine chlamydia trachomatis (Paavonen et al 1998) and toxoplasmosis during pregnancy screening (Lappalainen et al 1995) programmes have been evaluated. In addition the colorectal cancer screening programme has been preliminarily evaluated (Hristova et al 1996). An economic evaluation of the possible prostata cancer screening programme is being conducted at the moment by National Research and Development Centre for Welfare and Health. No evaluation has been conducted from the possible *Helicobacter pylori* screening programme in Finland.

3.2 Economic evaluation of ultrasound screening in pregnancy

The earliest economic evaluations of ultrasound screening in pregnancy reported only the rough unit cost of the screening ultrasound examination in pregnancy. None of the studies mentioned the principles underlying their estimates nor any detailed cost-accounting. Ewigman et al (1993), Persson et al (1983) and Temmermann et al (1991) reported the unit cost being \$23 to \$182 at 1990 prices. Bakketeig et al (1984) estimated the cost of screening at \$540/pregnancy at 1990 prices. This included two ultrasound screening examinations per woman, the increased use of hospital resources, loss of time and income for the attending women, and travel. The use of antenatal health services varies from more to less use of services in different studies (Persson et al 1983, Jörgensen 1992 and Backe et al 1990).

Two cost-effectiveness analyses have reported the cost per identified malformed fetus. DeVore (1994) reported the cost for identified malformed fetus as \$29 533 or \$115 575 depending on the diagnostic rate of the performers of the ultrasound screening (in tertiary centers vs, non-tertiary centers). The effectiveness data were based on the RADIUS- study, (RADIUS = Routine Antenatal Diagnostic Imaging With Ultrasound Study, 1987-1991 in USA, total number of diagnostic procedures performed 7281) which has been criticised of a low detection rate of malformed fetuses (Ewigman et al 1993). There was a significant difference in the number of malformed fetuses between

the tertiary and non-tertiary centers (6.8 vs. 1.7 per 1000, $p = 0.001$). The cost included the estimated health care cost of routine ultrasound examination, based on expert opinion. Non-health care or time costs were not included in the analysis.

Roberts et al (1998) reported the cost per target anomaly detected of one second trimester ultrasound being in the range 5000 pound sterling to 109 000 pound sterling. The objective of the study was to compare the cost-effectiveness of different programmes of routine antenatal ultrasound screening to detect four key fetal anomalies. The cheapest, but not the most effective, screening programme consisted of one second trimester ultrasound. The effectiveness and cost data were derived from expert opinion from the Royal College of Obstetricians and Gynaecologists, Working Party and secondary data from the literature. The authors concluded that the estimation model highlighted the weakness of available evidence and demonstrated the need for more information both about current practice and costs.

Waitzman et al (1998) have performed a preliminary benefit-to-cost analysis on a policy of routine ultrasound screening for fetal anomalies in the United States. The benefit to cost ratio ranged from 0.33 to 3. The cost and benefit estimation was based on the RADIUS study and European studies. The analysis focused on the cost savings associated with reducing congenital anomalies by terminating pregnancies. The direct costs of medical care, special education and development services were included, as were the indirect costs of reduction on future returns or productivity due to heightened morbidity or premature mortality. Vintzileos et al (2000) only used the RADIUS study data and reported the benefit to cost ratio to be between 1.35 and 1.70 in tertiary centers and between 0.40 and 0.74 in non-tertiary centers.

3.3 Economic evaluation of breast cancer screening

Breast cancer screening has been economically evaluated by several researchers. Brown et al (1999) have concluded in their review of the economic evaluation literature of breast cancer screening that a wide range of cost estimates exists across studies and that a lack of standardization exists across studies with regard to basic economic principles.

Cost-effectiveness analysis has been generally used, using the cost per life year saved (LYS) or the cost per quality-adjusted life year saved (QALY) as the main outcome measure in the analyses. However, studies of the cost-effectiveness of breast carcinoma screening have reported widely differing results, Table 1.

The early C/E studies which reported breast cancer screening to be cost-saving assumed very large differentials to exist between the lifetime treatment costs for late- and early-stage breast cancer, which the later studies have not confirmed (Moskowitz 1987, Holtzman 1990, Zavertnik et al 1992, Johnson 1988, Brown et al 1993).

The trial-based cost-effectiveness studies that exist have seldom been conducted from the societal perspective. i.e., the non-health care and time costs have been excluded. (U.S. Congress, Office of Technology Assessment 1987, Eddy 1989, Brown 1992, Okubo et al 1991, Kattlove et al 1995, Hall et al 1992, Lindfors et al 1995, van Ineveld et al 1993, Hristova et al 1996).The study of Tabar et al (1989) does not report the cost accounting methods at all.

Table 1-Previously published C/E ratios

	country	at prices (year)	cost/ LYS (US\$)
Gerard et al	Australia	89	14 800
Australian Health Ministry	Australia	89	7 897
Clarke et al	Scotland	89	69 817
Forrest	Great Britain	84	4 810
US Office of Technology	USA	87	34 600
Eddy	USA	84	83 830
Eddy	USA	84	21 717
Brown	USA	90	30 000
Forsyth	USA	1)	37 500
Okubo et al	Japan	1)	14 300
van der Maas et al	Netherlands	88	4 850
de Koning et al	Netherlands	90	3 825
Tabar et al	Sweden	1)	3 400
Kattlove et al	USA	93	8 280
Hall et al	Australia	89	5 321
Lindfors et al	general	1)	17 900
van Ineveld et al	UK	90	2 844
van Ineveld et al	Netherlands	90	3 350
van Ineveld et al	France	90	9 164
van Ineveld et al	Spain	90	15 326
Norum	Norway	1)	21 403
Hristova et al	Nordic countr.	1)	15 414

1) data not available

The estimated life-years saved and the base cost of screening are the essential basic elements of the cost-effectiveness analyses. The variation in the base cost of screening ranges from US\$18 to US\$86 in previously published studies (Table 1), reflecting the different cost structures in different countries, but also the different cost accounting methods.

The estimation of life-years saved is also variable in previously published studies. Most authors have assumed the breast cancer mortality reduction to be over 28% (Forrest 1987, U.S. Congress, Office of Technology Assessment 1987, van der Maas et al 1989, de Koning et al 1991, Lindfors et al 1995, van Ineveld et al 1993, Hristova et al 1996, Norum 1999) based mainly on the trial results from the Health Insurance Plan of Greater New York and the Swedish Trials (Shapiro et al 1988, Nyström et al 1993). The majority of published trials and meta-analyses have shown screening mammography to reduce breast cancer mortality in women above the age of 50 years (Shapiro et al 1988, Tabar et al 1985, Nyström et al 1993, Verbeek et al 1984, Kerlikowske et al 1995, Frisell et al 1997, Wald et al 1993). Interestingly, the Canadian, Edinburgh and Finnish trials report either no effect or a statistically insignificant effect (Miller et al 1992, Kerlikowske et al 1995, Hakama et al 1997, Roberts et al 1990, Alexander 1997). Meta-analysis by Gotzsche et al (2000) reported mammography screening not to reduce breast cancer mortality. The LYS depend crucially on whether early or late deaths are prevented and in this aspect most published trials have relatively short follow-up times.

Eddy (1989) and Okubo et al (1991) have estimated the number of life-years saved from an exceptional perspective. Eddy (1989) reported the marginal cost of LYS when mammography screening was added to physical breast examination. Okubo et al (1991) reported the cost-effectiveness of breast carcinoma screening for the age group 30-80 years for a period of 50 years, which was very different from any other screening programme or trial.

There is a broad agreement that in cost-effectiveness analyses all future costs and health consequences should be stated in terms of their current value, i.e., discounted. Interestingly, Gerard et al (1990), Forsyth (1988), Okubo et al (1991), Kattlove et al (1995), and Hristova et al (1996) did not discount the LYS.

No studies of the cost-effectiveness of nationwide population-based breast carcinoma screening have been published. It is noteworthy, that the clinical outcomes of the effectiveness that can be achieved are generally from small-scale trials involving a limited number of experts, persons examined, and areas. On a national scale, with hundreds of professional practitioners, it can be expected to be more difficult to attain uniform quality (de Koning 2000).

3.4 Economic evaluation of double-reading mammograms

The double-reading mammograms is a widely used criterion standard in breast cancer screening despite a lack of evidence of the cost-effectiveness of the second reading. The double-reading practice is reported to be more effective than single-reading practice (Anttinen et al 1993, Anderson et al 1994, Brown et al 1996, Ciatto et al 1995, Warren et al 1995, Blanks et al 1998). The method of double-reading has increased the number of breast cancers detected from 5% to 14 % in these studies.

Ciatto et al (1995) reported that double-reading detected cancers at an earlier stage than single-reading. This result did not reach statistical significance.

Ciatto et al (1995) reported double-reading to cause a marked increase in cost for each woman screened (8.5%) and a more limited increase in the cost for each cancer detected (3.5%) at the first screening. The costs included only direct health care costs of the base screening and the assessment procedures. The reporting of cost-accounting principles was limited in this study. Cost – effectiveness was not estimated.

Only one publication on this topic takes comprehensively into account both effects and costs (Brown et al 1996). It reports consensus double-reading to be more effective and less costly in Britain than a single reading policy. The effectiveness data of this study were from an observational, non-randomised trial of 34 000 attenders. A sample of 132 attenders provided data for private costs. Costs were estimated from a societal perspective. Brown et al (1996) compared three reporting policies: single reading, consensus double-reading, and non-consensus double-reading. Consensus double-reading costs less than single reading (saving £4853 per 10 000 women screened). Non-consensus double-reading costs more than a single reading (difference £19 259 per 10 000 women screened).

3.5 Economic evaluation of *Helicobacter pylori* screening

The previous published studies have estimated the financial implications of screening for *H pylori* in one subpopulation, the dyspeptic patients or from the viewpoint of one *H pylori* related disease, gastric cancer.

Briggs et al (1996) have estimated the cost-effectiveness of identifying patients with *H pylori* and peptic ulcer and of providing eradication treatment compared with conventional empirical *H sub 2*-antagonists treatment. They report the screening strategy to result in savings after eight years from the screening. Only health care costs were included in the Markov model.

Fendrick et al (1999) and Parsonnet et al (1996) reported, that screening and treatment for *H pylori* is potentially cost-effective in the prevention of gastric cancer. These analyses do not take into account the benefits of prevention of other *H pylori*-related diseases. Cost-effectiveness were 6264 US\$, at 1996 prices (Fendrick et al 1999) and 25 000 US\$, at 1992 prices (Parsonnet et al 1996) per life year saved in the base case analyses. The screening was conducted at age 40 (Fendrick et al 1999) and 50 (Parsonnet et al 1996) years, prevalence of *H pylori* infection was estimated as 40% , relative risk of cancer in persons with *H pylori* infection as 3.6. *H pylori* was estimated to prevent 100% (Parsonnet et al 1996) and 30% (Fendrick et al 1999) of distal cancers. Distal cancers were assumed to comprise 50-80% (Fendrick et al 1999) and 60% (Parsonnet et al 1996) of all gastric cancers. Only health care costs were included in these Markov models. Fendrick et al (1999) did not include treatment costs for gastric cancer or medical costs for unrelated illness in added life-years in

their baseline analysis. Parsonnet et al (1996) did not include screening office visit cost, but did include medical costs for unrelated illness in added life-years.

IV AIMS OF THE STUDY

The purpose of the present study was to conduct an economic evaluation of two nationwide Finnish screening programmes, and one pilot Finnish screening programme aimed at nationwide use. The reason for choosing these programmes for analysis was, that data had become available for conducting economic evaluation.

The specific aims of this study were:

- 1 To evaluate the cost-effectiveness of one-stage second trimester ultrasound screening in pregnancy in Finland.
- 2 To evaluate the cost-effectiveness of nationwide breast carcinoma screening in Finland, 1987-1992
- 3 To evaluate the cost-effectiveness of double-reading mammograms in Finland.
- 4 To carry out a cost-benefit analysis of *Helicobacter pylori* screening in Finland.

V MATERIAL AND METHODS

5.1 Effectiveness data (I, II,III)

The economic analysis of the one-stage ultrasound screening in pregnancy (I), mammography screening (II) and double-reading mammograms (III) was carried out as a cost-effectiveness analysis.

The ultrasound effectiveness data (I) came from population-based clinical Helsinki Ultrasound Trial (Saari-Kemppainen et al 1990, 1994, 1995), which took place from 1986 to 1988, and incorporated 9310 women; 95% of all pregnant women in the Helsinki metropolitan area. They were randomly allocated to a screening group or a control group. Only four women dropped out during the follow-up.

Ultrasound examinations for the screening group took place between the 16th and 20th gestational weeks at the Helsinki City Hospital and the University Hospital, and were performed by ultrasound-trained nurses. Repeat ultrasound was necessary for 6.5% of the women, some of whom also needed further examinations and procedures performed by obstetricians. The data on ultrasound findings were collected on pre-structured forms. At the University Hospital, more time than normal was allocated to the study screening examinations. Apart from the study ultrasound examination, the screening group and control groups received the same antenatal care, which may have included ultrasonography for other than screening purposes at the mother's own doctor's discretion.

The effectiveness of the ultrasound screening was measured in terms of perinatal mortality in the screening group compared to the control group. Data on perinatal mortality were collected from hospital records and health-center records of postnatal health-center visits.

The effectiveness of mammography screening (II) was measured by life-years saved (LYS) due to screening. Finland has had a nationwide population-based breast cancer screening programme since 1987, which covers women from 50-59 and can be continued up to age 64. The women are screened every second year. The programme started gradually with cohorts born in even years, with those born in odd years serving as controls, during the first few years of operation.

In the study to the end of 1992 the annual original data (Hakama et al 1997) were used for cumulative breast cancer mortality (excluding those cancers diagnosed before the screening) in the screening and control groups in the national programme. The number of invitees was 89,893, number of participants 76,389 and number of controls

68,862. The screenees were invited to join the program in 1987-89. They were born in 1927-1939 and screened by the Finnish Cancer Society. The controls were the women, who were not invited to the screening, matched by age and municipality of residence. The follow-up was extended to the end of 1992. Breast cancer cases diagnosed at screening, interval cancers, cancers diagnosed among the women invited but not screened, cancers diagnosed in the control cohorts, and deaths from breast cancer was identified by linkage to the Finnish Cancer Registry (Hakama et al 1997).

From the beginning of 1993 an estimation model was used, i.e., the total LYS were estimated by adding the LYS in 1987-1992 to the LYS estimated by the model from 1993 onwards. The LYS were estimated to the year 2020, i.e., when the discounted annual LYS would be under 5 years, and the average age of the screening group 87 years.

Of the breast cancer cases diagnosed from 1987 to 1992, assuming no screening, the number succumbing annually between 1993 and 2020 was estimated based on the age-specific 1980-84 national cumulative breast cancer survival rates (Finnish Cancer Registry) for the 12 years following diagnosis. After 12 years, national mortality rates at age of the patients were applied to the patient population with more than 12 years of survival. The survival rates were multiplied by 0.985, indicating persistent excess relative mortality due to breast cancer among the patients. In addition, the cumulative relative mortality was corrected by a constant multiple of 0.8 due to an assumed 20% improvement in breast cancer survival from 1993 to 2020 compared to the period 1980 to 1984.

The LYS were estimated under three optional models. In the base case model screening was assumed to lower cumulative relative (i.e. breast cancer) mortality by 0% in the first year after diagnosis, by 12% in the second year and by 24% in the third and fourth years. In the fifth year, the cumulative relative mortality reduction, was assumed to equal the reduction of the fourth year plus 12% of the marginal, (i.e. additional) relative, mortality. After the fifth year, cumulative relative mortality reduction was assumed to remain constant, i.e., annual mortality in the screening and the no-screening groups was assumed to remain the same. After the fifth year the cumulative relative mortality reduction was not assumed to increase. (Shapiro et al 1988, Nyström et al 1993, Hakama et al 1997).

In the second case model, the screening effect was estimated to equal that of the first model during the first four years after diagnosis. In the fifth year, the cumulative relative mortality reduction was assumed to be the reduction of the fourth year plus 18% of the marginal relative mortality. In the sixth year, cumulative relative mortality reduction was assumed to be the reduction of the fifth year plus 12% of the marginal relative mortality. In the seventh year, the cumulative relative mortality reduction was assumed to be the reduction of the sixth year plus 6% of marginal relative mortality. After the seventh year, annual mortality in the screening and the no-screening groups was assumed to remain the same.

In the third case model, annual mortality was assumed to remain the same in the screening and no-screening groups after the end of the screening in 1992.

In all three models the all-cause life-years saved were obtained by annually multiplying cumulative relative breast cancer-mortality reduction by the age-specific population mortality.

The double-reading effectiveness material (II) consists of 95 423 mammographs performed during 1990-1995 in three screening centers (Pori, Jyväskylä, and Kuopio) of the Finnish Cancer Society, representing about 15% of the total screening activity in Finland during that period. The participation rate was 91%.

The material is retrospective, measuring the effectiveness of the actually used double reading strategy. All mammograms are read by two radiologists, chronologically the first and second reader. However, in practice, the second reader had access to the reading result of the first, i.e., the double reading was not all blind. In addition to the double reading, some centres have had consensus reading for some of the mammograms. Consensus reading was performed by one senior radiologist or by a group of two to four radiologists.

The centres recorded the readings of the first and the second reader on a grading scale of: 1 normal, 2 benign, 3 mild malignant suspected, 4 strong malignant suspected, 5 malignant. Grades 3-5 mean that the reader considers further assessment useful. The number of further assessment procedures, the number of surgeries and the diagnosis are also recorded.

We compared two strategies for reporting within a screening program:

- 1) Single-reading strategy with possible consensus reading, i.e., assuming only one reader performed the original reading.
- 2) Double-reading strategy with possible consensus reading, i.e., the strategy actually used.

The main indicator of effectiveness was the actual number of cancers detected, based on their histological diagnosis. This number of cancers detected by single-reading strategy was estimated assuming that all the cancers originally graded 3-5 by the first reader would have been detected by the single reading, but none of graded 1-2 by the first reader would have been thus detected. The effectiveness of the double reading was the difference between number of cancers detected in the double-reading strategy and that for the single reading. The chi-square test served to compare incremental effectiveness of different screening centers statistically. Due to the incomplete data on interval cancers within the Finnish national screening program, the definition of effectiveness did not include the potential where the initial reader read 3-5 and the

second reader 1-2 with subsequent consensus of 1-2 suggesting the first reader was an possible “over” read.

Differences between double and single reading strategies were further assessed by the mean size of lesions in mammography, the mean size of lesions in microscopy and stage distribution. Data for these comparisons came from two screening centers (Kuopio and Jyväskylä). Student’s t-test and the chi-square test served in analysis of statistical differences.

5.2 Cost data (I, II, III)

Costs were evaluated from the societal perspective, i.e., including health-care, non-health care and time costs. The ultrasound study (I) was conducted a couple of years before the breast cancer(I) and double reading mammograms studies (III). The generally used cost terminology was in a change during that time, which led to variability in the use of cost terms among the studies. In the ultrasound screening study terms direct and indirect costs are used. Direct costs were costs which were directly caused by the health program. Indirect costs were lost working time costs and travel costs. Positive costs were actual realised costs, and negative costs represented savings due to lower use of health services. When representing the results, the following terms were used: gross cost = only the positive costs included, net cost 1 to 3, the positive costs included, but negative costs, i.e. savings to different degrees included. In the mammography screening study (II) and double reading mammograms study (III) the costs were divided according to Gold et al (1996) into health care, non-health care and time costs. Health care costs included also the health care savings caused by the screening programme. Non-health care costs were travel costs.

In the ultrasound screening in pregnancy (I) the main cost-causing events were the base cost of screening, the ultrasound-induced advanced examinations, and the savings, i.e., negative costs due to fewer inpatient and outpatient visits and fewer other ultrasounds. Later we also obtained data of the additional use of other health services, which represent part of the total cost. These data were collected from every woman participating in the trial. The differences between the comparison groups were evaluated by Student’s t-test.

Cost-accounting data were collected via a questionnaire for all attenders, and were later completed by a postal questionnaire to 534 of the screened women. The sample was a stratified random sample to equal sample size by health center. The response rate was 76%. The responding and non-responding groups were fairly similar, measured by income, age and distance to health center. Further cost accounting sources were internal accounting and health-market data and statistics.

In the mammography screening (II) and double-reading mammograms (III) the main cost-causing events were: 1 base costs of screening, 2 costs of recall assessment, 3 surgery costs for false positives, 4 savings due to early treatment of breast cancer. In the mammography screening study we also estimated the costs of treatment caused by over-diagnosis. In double-reading costs we included all the relevant incremental costs associated with the double-reading strategy compared to the single-reading.

The first mammography represented the base health care cost of screening. In Finland a two-view projection is used. The costs were derived from the 1995 internal accounts from the four main screening organizations in Finland. Adjustments were made to standardize the different accounting methods. The mean national base health care cost of screening and assessment was calculated by weighting the results with the market shares of the organizations.

The base non-health care cost of screening was for transportation. The time cost was the opportunity cost of time actually spent on the mammography, as well as the transportation and waiting times. The cost data were collected by questionnaire from 1400 attendees at the screening examination, a stratified random sample to equal sample size by the national statistical grouping into residents by municipalities (Statistics Finland 1995). The questionnaire was distributed in 10 randomly selected screening units in Finland from April 1996 to June 1996. We received 1294 returned questionnaires (92%). All the cost estimates were aggregated from individual data. When any piece of information was missing, it was replaced by the sample average. The opportunity cost of time was regarded as zero for women who didn't use working time for the examination. For those women, who were employed at the time of examination and who used working time for the examination, the opportunity cost of time was the time spent multiplied by the hourly wage rate (including employer's social expenses; national average 40.7% of the salary).

The health care cost of recall assessment consisted of the repeated or detailed mammograms, ultrasound and ultrasound + fine needle biopsy costs. The proportion of women further examined and the proportion of fine needle biopsies came from a Finnish study (Pamilo et al 1995), showing the equivalent national figures from all screening units in Finland during this period. The proportion of repeated mammographs and ultrasounds of women further examined in total was assumed to be the same as during 1990-95 in Kuopio and Jyväskylä, two screening units of the Cancer Society of Finland. Under these assumptions, 3.5% of the screened women were recalled for further assessment, 2.7% were remammographed, 2.3% underwent ultrasound, 0.9% fine-needle biopsy and 0.7% surgery. The non- health care cost of transportation and the time cost were assumed to be 1.5 times higher than for the base screening due to the longer distances to the assessment locations.

In the double-reading study (II) the number of incremental further assessment procedures were recorded in each screening centres. The mean national cost served as basis for the cost estimate.

The health care costs for surgery on the false positives were from two major hospitals for breast cancer surgery in Finland: Helsinki City Hospital and Kuopio University Hospital. The transportation cost for the surgery for positives was assumed to be twice as high as in the mammography screening. Evaluation of time cost was based on the results of the cost survey of the screening attendees and the data on average sick leave caused by the operation. It was assumed that all employed women used working time for the procedure and the following sick leave. The surgery cost of the true positives is included in the saving due to early treatment.

The saving due to early treatment was based on data on the difference in stage distribution with and without screening. The data on the national stage distribution from 1980-84 and 1990-94 came from the Finnish Cancer Registry, and were defined as localized or non-localized. The period 1980-84 represented the no-screening situation and 1990-94 the screening situation. The treatment cost for the localized and non-localized cancers came from a Finnish study (Parvinen et al 1996), of costs of all breast cancers in the area of Turku in 1981 (30 non-localized and 20 localized breast cancers), for which health-resource use was followed up for nine years or the patient's death. Outpatient visits and inpatient days for localized and non-localized cancer are based on a study of 535 patients diagnosed in 1977-80 in the Tampere area and followed up for five years after diagnosis or until death (Holli et al 1996). Transportation time was estimated as twice as high as in the base mammography screening, since the treatment mainly took place in the five University Hospitals. The time cost was calculated assuming that all the employed women used working time for the treatment.

Cost of treatment caused by over-diagnosis is due to the possible over-diagnosis of carcinoma in situ in the cancers detected by screening. The natural history of carcinoma in situ is unknown, but it seems to be a high-risk marker rather than an invariable precursor of invasive disease (Wright et al 1995, Ernster et al 1996, Page et al 1982, Rosen et al 1982). In the age group 50-59 the national number of new breast carcinoma in situ (ductal and non-ductal) and new breast carcinoma in total were obtained from the Finnish Cancer Registry for the period 1980-84 and 1990-94. The 1990-94 proportions were applied to the screening group. The cost of treating 72% (Page et al 1982) of those new intraductal breast cancers and 62% (Rosen et al 1978) of those new non-ductal breast cancers in situ which exceeded the proportion of equivalent carcinomas in situ to carcinomas in total estimated from the 1980-84 data, was assumed to be the cost of over-treatment cost of breast carcinoma due to screening.

5.3 Measure of cost-effectiveness (I, II, III)

The cost-effectiveness of ultrasound screening in pregnancy (I) was estimated by dividing cost difference by the difference in perinatal deaths between the screening

and control groups. The main outcome measure was the cost of avoiding one perinatal death.

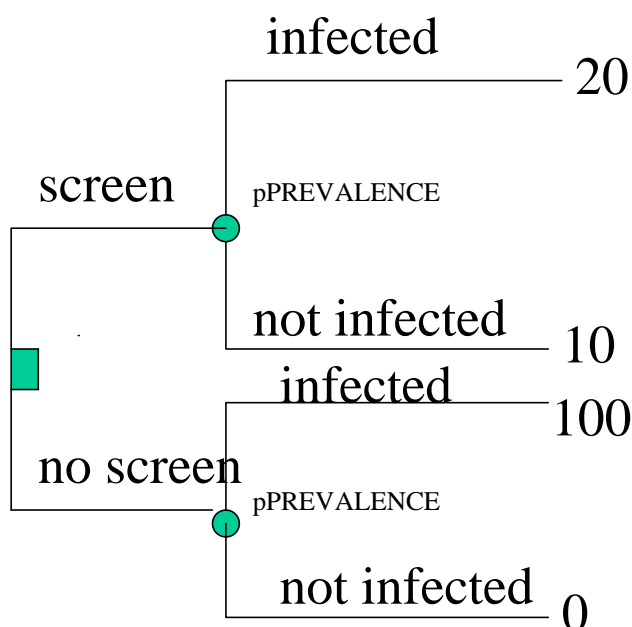
The cost-effectiveness of mammography screening (II) was estimated by dividing the costs of screening in the period 1987-92 by the total estimated number of life-years saved in the period 1987-2020 according to the three sets of assumptions described in models 1 to 3.

The main outcome measure of the cost-effectiveness of double-reading mammograms (III) was the incremental cost per additional cancer found due to the practice of double reading.

5.4 Costs and benefits in cost-benefit analysis (IV)

The economic evaluation of *Helicobacter pylori* screening was carried out as a cost-benefit analysis using a computer-based decision tree (SMLTree). A full cost-benefit analysis measures all costs and consequences in dollars. However, in practice cost-benefit analyses often amount to a comparison of those costs and benefits, that can be easily expressed in money terms (Drummond et al 1997). Also in our cost-benefit analysis the scope was restricted. Only health care costs were included in the analysis, i.e. non- health care costs and time costs were not included. Health care costs did also include the possible health care savings, which thus represented the benefit of the screening. We did not put monetary value on changes in health state, e.g. death due to gastric cancer. Thus, some researchers would classify the present analysis as cost analysis, not cost-benefit analysis.

A decision tree is composed of three types of nodes. Decision node branches correspond to strategies that the decision maker has control over, i.e. to screen or not to screen. Chance node branches correspond to events over which the decision maker has no control, but can predict with certain probabilities, e.g., the lifetime probability of developing gastric cancer. Terminal nodes correspond to final outcome states. They contain utility expressions, i.e. monetary values, which describe the value of the state. When a chance node is evaluated, the probability of each branch is multiplied by the value of the node at the end of the branch and all such products are summed up to form an expectation (Hollenberg 1993).



The primary decision analysis compared two intervention strategies. 1) screen for H pylori and treat those individuals with a positive test, and 2) do not screen for H pylori and test and treat H pylori only if related clinical symptoms appear. The model estimated the discounted H pylori-related accumulated health care costs from screening age to death in both strategies. The age at screening was 15-45 years in the base case and 15, 30 and 45 years in the sensitivity analyses.

The main outcome measure is the incremental health care cost per case, i.e., person invited for screening, in the screening alternative compared to no-screening alternative. In addition, the incremental cost per treated H pylori infection due to screening is presented.

The participation, visit and compliance rates, prevalence of H pylori antibodies, effectiveness of H pylori treatment (i.e. eradication rate), and sensitivity and specificity of the H pylori serological screening test were obtained from the Vammala H pylori screen and treat project (Salomaa et al 1998). The population-based screen and treat project started in 1996 in Vammala, a semi-urban community in southwest Finland. In 1996, all subjects aged 15 to 40 years, and in 1997 and 1998, all subjects aged 15 and 45 years, living in Vammala, were invited by mail to participate in the project. Serum IgG and IgA antibodies to H pylori were determined by enzyme immunoassay, the lower limits of raised titres were 700 for IgG and 70 for IgA antibodies. Within these limits, the sensitivity and specificity of the tests were 94% and 93%, respectively (Kosunen et al 1997). The minimum and maximum values for sensitivity analyses were obtained from published studies (van Der Ende et al 1999, Meijer et al 1997, Xia et al 2000). All antibody-positive subjects were invited to visit

a GP and offered eradication therapy . The first-line therapy was amoxycillin 1000 mg x 2, metronidazole 400 mg x 3 and lanzoprazole 30 mg x 2 for one week. Serological follow-up occurred 6 months later. Treatment failures were offered a second eradication therapy by phone. Second treatment failures were offered gastroscopy, including antibiotic sensitivity testing, which served as a basis for the third individual treatment regimen. In 1996-1998, 5288 subjects entered the project. These data were used in the current analysis, with the exception that the cost of the one-week therapy was replaced with the cost of the currently recommended therapy; i.e., amoxycillin 1000 mg x 2, clarithromycin 500 mg x 2 and omeprazol 20 mg x 2 for one week. To simplify the model, no third treatment failure was assumed to exist, and the sensitivity and specificity of the retests were assumed to equal to 100%.

Participation rate was defined as the number of patients who gave serum samples / the number of invited patients. Visit rate was the number of patients who visited the GP / the number of patients who had positive screening samples. Compliance rate was the number of patients who took the prescribed drugs and gave a follow-up sample 6 months later / the number of positive patients who visited the GP. The second and third compliance rates were the number of patients who took the prescribed drugs and gave a follow-up sample 6 months later / the number of patients who had positive follow-up samples.

At screening, the prevalence of Hp antibodies, indicating infection was 13%, participation rate 76%, visit rate 87% and compliance rate 91% in 15 to 45-year-olds. The effectiveness of the first treatment was 81%. Neither re-infection after cure nor spontaneous eradication in the absence of therapy were included in the model (Table 2).

Table 2. Probability estimates for H pylori screening-related variables in Decision Tree Analysis.

Variable	Probabilities					
	Base		Alternative			
	Case		Models			
		Age at screening(years)		Min.	Max.	
		15	30	45	Value	Value
Participation rate	0,76	0,91	0,74	0,70	0,61	0,90
Prevalence	0,13	0,03	0,13	0,29	0,10	0,16
Screening test (serology) sensitivity	0,97				0,90	0,98
Visit rate	0,87	0,75	0,87	0,79	0,70	0,90
First Compliance rate 1)	0,86	0,78	0,85	0,91	0,69	0,90
Effectiveness of first treatment (one week)	0,81				0,65	0,90
Second compliance rate 2)	0,47					
Effectiveness of second treatment	0,82					
Third compliance rate 2)	0,50					
Screening test specificity	0,93				0,90	0,98

VS = Vammala Helicobacter pylori screen and treat project

* maximum value estimated at 0.90

1) number of patients who took prescribed drugs and gave follow-up sample 6 months later / number of Hp- positive patients who visited GP

2)number of patients who took prescribed drugs and gave follow-up sample 6 months later/ number of patients who had positive follow-up sample

Probability estimates for H pylori – related disease variables were obtained from published reports and national statistics (Table 3, Forman 1995, Forman 1994, The

Eurogast Study Group 1993, Graham 1997, Parsonnet 1998, Kuipers 1999, Kokkola et al 1996, Calam 1998, Ihämäki et al 1979, Sipponen et al 1990, Cullen et al 1993, Dixon et al 1992, Sipponen 1995, Dew 1987, Varis 1988, Katschinski et al 1994, van Der Hulst et al 1997, Elashoff et al 1983, Danesh et al 2000, Miwa et al 2000, Moayyedi et al 1999, Greeberg et al 1999, Talley et al 1999, Gilvarry et al 1997, McColl et al 1998, Peeters 1998). *Helicobacter pylori* infection was regarded as a risk factor for gastric cancer, gastric premalignant conditions, duodenal ulcer, gastric ulcer and in certain analyses for dyspepsia.

Table 3. Probability estimates for H pylori-related disease variables in Decision Tree Analysis.

Variable	Probabilities			
	Base		Alternative	
	Case		Models	
		Min.	Max.	
		Value	Value	
Cancer, Lifetime probability of gastric cancer in general population a)	0,004			
Cancer, Hp-related gastric cancers of all gastric cancers	0,73			
Cancer, Lifetime probability of Hp-related gastric cancer, if Hp+ b)	0,0225	0,005	0,03	
PL, Lifetime probability of diagnosed Hp-related PL, if Hp+	0,0898	0,02	0,12	
PL, Probability of operational treatment	0,11			
DU, Lifetime probability of DU in general population	0,1			
DU, Proportion of DU Hp+	0,95			
DU, Hp prevalence in population, where DU probability studied	0,8			
DU, Lifetime probability of Hp-related DU, if Hp+ c)	0,12	0,11	0,22	
DU, Diagnosed in emergency care	0,05			
DU, Probability of survival, when DU diagnosed in emergency care	0,92			
DU, Probability of bleeding complications	0,0042			
GU, Lifetime probability of GU in general population	0,035			
GU, Proportion of GU Hp+	0,8			
GU, Hp prevalence in population, where GU probability studied	0,8			
GU, Lifetime probability of Hp-related GU, if Hp+ d)	0,028	0,02	0,04	
GU, Diagnosed in emergency care	0,05			
GU, Probability of survival, when GU diagnosed in emergency care	0,92			
GU, Probability of bleeding complications	0,0021			
Dyspepsia, Lifetime probability of Hp-related funct. dyspepsia, if Hp+	0,05	0	0,1	
Dyspepsia, Helicotest UBT sensitivity	0,98			
Dyspepsia, Compliance	0,91	0,73	0,95	

a) Finnish Cancer Registry

b) = $(1/\text{prevalence}) \times (\text{pcancer} \times \text{proport.H pylori} +)$

c) = $0,1 \times 0,95 \times (1/0,8)$

d) = $0,035 \times 0,8 \times 0,8 \times (1/0,8)$

PL = Premalignant lesions

Hp = *Helicobacter pylori*

DU = Duodenal ulcer

GU = Gastric ulcer

VS = Vammala *Helicobacter pylori* screen and treat project

The observed sex-adjusted gastric cancer incidence rates by birth cohort were obtained from the Finnish Cancer Registry. The gastric cancer statistics include all cancers whose primary site is the stomach, e.g., gastric lymphomas, which are also strongly associated with *H pylori* (Wotherspoon et al 1991). The cohort-specific risks have decreased quite steadily over time, with the risk between successive cohorts appearing to be fairly constant (Hakulinen et al 1986). We assumed that the ratio of the risks for successive birth-year cohorts to be fairly constant for the 1968 birth cohort (aged 30 years in 1998). The lifetime risk of developing gastric cancer in the

general 1968 birth cohort was obtained by multiplying the estimated age-specific gastric cancer incidence rates by the age-specific population mortality (Statistics Finland 1999) and then summing up these figures. The lifetime probability of gastric cancer was estimated to be 0.4% in the 30-year-old general population. We assumed that 73% of all gastric cancers are caused by H pylori infection. (Forman 1995, Forman 1994, Eurogast Study Group 1993). The lifetime probability of infected 1968 birth cohort to develop H pylori caused gastric cancer was estimated at 2.25 % $(1/\text{prevalence}) \times (\text{probability of cancer in the general 1968 birth cohort}) \times (\text{proportion of cancers caused by H pylori})$. In the sensitivity analysis, minimum and maximum values of 0.5% and 3.0% were used (Parsonnet 1998, Graham 1997). (Table 3)

The lifetime probability of diagnosed H pylori-related premalignant conditions in infected patients was estimated at 8.98%. Kuipers (1999) reported an 8-fold risk of developing gastric premalignant lesion as compared with gastric cancer. However, his baseline risk estimate for gastric cancer was only 1%. We conservatively assumed a 4-fold risk of developing gastric premalignant conditions as compared with gastric cancer (Table 3).

The lifetime probabilities of H pylori related duodenal or gastric ulcers in infected patients were estimated as follows: $(\text{lifetime probability of duodenal/gastric ulcers in general population}) \times (\text{proportion of duodenal/gastric ulcers caused by H pylori}) \times (1/\text{H pylori prevalence in the age cohort for which duodenal/gastric ulcer probability used})$. The estimated lifetime probabilities of H pylori related duodenal and gastric ulcers in infected patients were 12% $(0.1 \times 0.95 \times (1/(0.8)))$ and 2.8% $(0.035 \times 0.8 \times (1/(0.8)))$ in the baseline case (Kokkola et al 1996, Calam 1998, Ihämäki et al 1979, Cullen et al 1993, Dixon et al 1992) (Table 3). The probability of ulcer bleeding after H pylori diagnosis and treatment was estimated to be as low as 0.4%, because there is strong evidence for the benefits associated with the current clinical practice of eradication H pylori in patients with a confirmed ulcer (van der Hulst et al 1997, Elashoff et al 1983).

The lifetime probability of H pylori related diagnosed dyspepsia in infected individuals is controversial (Danesh et al 2000, Miwa et al 2000, Moayeddi et al 1999, Greenberg et al 1999, Talley et al 1999, Gilvarry et al 1999, McColl et al 1998). We assumed the lifetime probability of H pylori related diagnosed dyspepsia in infected patients in the baseline at 5%. (Table 3).

The life-time probability of Helicobacter pylori-related morbidity, treatments and mortality were estimated to be equal for the screening group non-participants and the no-screening group.

Only health care costs were included in the model. Unit treatment cost data were obtained from Helsinki University Hospital, Uusimaa Health District (Helsinki University Central Hospital 1999, Uusimaa Health District 1999) and from in-house cost data of Helsinki City Department of Health. The drug costs were at retail market

prices (Lääketietokeskus 1999). Treatment-associated adverse events and their treatment were included in the model.

The aggregated health care cost of gastric cancer in the baseline case was estimated as US\$ 13 413. This included costs before and after diagnosis. The average number of different cost events prior to diagnosis was obtained from a published article (Suvakovic et al 1997) and complemented by expert opinion. The source of gastric cancer treatment cost after diagnosis was the National Research and Development Center for Welfare and Health. All new gastric cancer patients (n= 467) diagnosed in 1992-93 in the Helsinki University Hospital district were included in the cost analysis. Patient-specific numbers of in-patient treatment periods in regional, central, university or other hospitals were obtained from the national statistics. In-patient treatment periods in the 5 years following cancer diagnosis were included when gastric cancer was recorded as one of the main diagnosis for treatment. The average unit costs for in-patient treatment periods and outpatient visits in different hospital types were obtained from hospitals' in-house cost data. This was also the source for the average number of outpatient visits per year for a gastric cancer patient. The average number of patients using outpatient services was estimated from the annual relative survival (Dickman et al 1999) and all-cause survival rates (Statistics Finland 1999) (Table 4).

The costs of gastric premalignant conditions, duodenal and gastric ulcers and dyspepsia were the resource use costs. The relevant cost-causing events were estimated partly based on expert opinion and partly on published articles. All H pylori infections were estimated to be diagnosed in patients with confirmed gastric premalignant lesions, duodenal or gastric ulcers. H pylori serology tests in addition to H pylori gastroscopy specimen tests were estimated to have been taken in 8% of cases (Laine et al 1997). Compliance with H pylori eradication treatment was assessed to be 100%, in patients with gastric premalignant lesions, duodenal or gastric ulcers. Compliance was estimated to equal the screening (91%), if the patient had been diagnosed with dyspepsia. The effectiveness of H pylori treatment was estimated to equal the effectiveness rate observed at screening. (Table 4).

The cost-causing events of gastric premalignant lesions included events prior to diagnosis, diagnosis, eradication and follow-up. Follow-up consisted of 3 gastroscopies and specialist visits (Kokkola et al 1996), taking into account the all-cause mortality. The main cost-causing events of duodenal and gastric ulcers were events prior to diagnosis, diagnosis, H pylori eradication treatment and possible bleeding complications. The cost-causing events of dyspepsia included events prior to diagnosis, immediate or delayed diagnosis and eradication. (Table 4).

Table 4. Selected undiscounted aggregated costs (US\$ 1998/99) in Decision Tree Analysis.

Cost-causing event	US\$
Screening, participate, infected, test sensitive, visit, compliant, 1st treatment effective	232
Screening, participate, infected, test sensitive, visit, compliant, 2nd treatment effective	308
Screening, participate, infected, test sensitive, visit, compliant, 3rd treatment effective	722
Screening, participate, not infected, test specific	29
Screening, participate, not infected, test specific, visit, compliant	232
Gastric cancer, treatments and examinations prior to diagnosis a)	1227
Gastric cancer, treatment	13434
Premalignant lesion, treatments and examinations prior to the diagnosis a)	713
Premalignant lesion, treatment, surgery	4841
Premalignant lesion, treatment, no surgery	2999
DU, treatments and examinations prior to diagnosis a)	478
DU, emergency diagnosis, survival, later bleeding	7154
DU, emergency diagnosis, survival, no later bleeding	2498
DU, emergency diagnosis, no survival	9340
DU, no emergency diagnosis, later bleeding	5714
DU, no emergency diagnosis, no later bleeding	1058
GU, treatments and examinations prior to diagnosis a)	709
GU, emergency diagnosis, survival, later bleeding	7720
Dyspepsia, treatments and examinations prior to diagnosis a)	209
Dyspepsia, test sensitive, compliant	800
Dyspepsia, test sensitive, not compliant	937
Dyspepsia, test not sensitive, later diagnosis and treatment	989

Hp = *Helicobacter pylori*

DU = Duodenal ulcer

GU = Gastric ulcer

a) costs prior to diagnosis are included in aggregate costs

5.5 Valuation and discounting (I, II, III, IV)

In the ultrasound study all cost data were estimated at the 1990 price level, in the breast cancer screening and double-reading mammograms at the 1995 price level and in the H pylori study at the 1998/99 price level using indices. FIM values were converted to US\$ by the official exchange rates, i.e. US\$1.00 = FIM 3.85 in 1990, US\$ 1.00 = FIM 4.37 in 1995 and US\$1.00 = FIM 5.54 in 1999.

In the ultrasound study no discounting was needed, because the time period was the length of one pregnancy. In the breast cancer screening, double-reading mammograms and H pylori screening studies a discount rate of 3% was used (Gold et al 1996).

For the timing of the cost-causing events the data on attendance and difference in breast cancer mortality by calendar year came from the Finnish Cancer Registry. The number of screened women was recorded annually for the years 1987-89. The number of screened women in 1990-92 is based on the estimate that 90% of the women who once have entered the screening program, will continue in the program and will be screened every second year until 1992. The national mean participation rate has been 88.6% in Finland for the years 1987-94 (Pamilo et al 1995). The assessment and surgery were assumed to take place in the same calendar year as the screening. The savings were assumed to be equally distributed over the five years following

diagnosis, because more specific data were unavailable. Over-treatment costs were equally distributed over the two years following the diagnosis, because carcinoma in situ is a localised carcinoma, usually treated immediately after diagnosis.

The average age at diagnosis of H pylori caused disease was obtained from published articles. The age at diagnosis was 69 years in gastric cancer (Dickman et al 1999), 63 years in gastric premalignant lesions (Ihamäki et al 1979), 51 years in duodenal ulcers and 59 years in gastric ulcers (Kekki et al 1990). The average age of 45 years at diagnosis of dyspepsia was based on expert opinion.

5.6 Sensitivity analysis (I, II, III, IV)

In the ultrasound study (I) the costs and cost-effectiveness were calculated under five different model variations:

Table 5. Cost calculation model variations in the ultrasound screening study

cost type	gross cost	Net cost 1	net cost 2	net cost 3	net cost 4
screening induced costs	x	X	x	x	x
savings from other use of antenatal outpatient services		x	x	x	x
savings from other use of antenatal inpatient services			x	x	x
savings from other use of other ultrasounds				x	x
Costs of other use of antenatal maternal health center					x

Univariate sensitivity analysis was applied to proportionally large or uncertain cost items, and the changes used were +/- 5% and +/- 10%.

The cost-effectiveness of mammography screening (II) was estimated under different model assumptions: I base case, II surgery cost only for clinically evaluated, unnecessary false positives (Jahkola et al 1991), III no over-treatment cost, IV time cost for all attendees, based on the average salary of the sample, V no transportation and time cost, VI discount rate 5%, VII no over-treatment, transportation and time

cost, discount rate 5%, VIII difference in breast cancer survival constant after 8 years from diagnosis in screening and no-screening groups, IX difference in breast cancer survival constant after end of screening (1992) in screening and no-screening groups. In the univariate sensitivity analysis the base case model parameters were changed by +/-5%.

The incremental cost-effectiveness ratios of double-reading mammograms (III) were estimated for each of the three screening centers separately. The ratio was also estimated solely for invasive cancers, excluding the carcinoma in situ cases. This estimation was based on the sub-sample of two screening centers, assuming that the incidence of carcinoma in situ in the third center (Pori) was the weighed average of the sub-sample of the two centers (Kuopio and Jyväskylä). In addition we calculated the incremental C/E ratio when the base case model parameters were changed by +/- 10 %, i.e., univariate sensitivity analysis was performed.

In the *Helicobacter pylori* study univariate sensitivity analyses were performed by changing the values of selected key baseline variables. The screening - related baseline probabilities were changed by +/- 20%, notwithstanding the probabilities of participation, visit, compliance and effectiveness of treatment rates, which were not estimated to exceed 90%. The analysed ranges for screening test sensitivity and specificity were 90-98%. The *H pylori*-related disease probability ranges were obtained from published articles (Tables 2,3).

The base cost of screening, i.e., the costs of the screening test and GP visit, was changed by +/- 20% in the sensitivity analysis, as well as the aggregated costs of *H pylori*-related diseases.

In multivariate sensitivity analyses, we estimated the cost-benefit of *H pylori* screening when the age at screening was 15, 30 or 45 years. The simultaneously adjusted age-specific variables were participation, visit and compliance rates, prevalence and all the discounted cost variables. The age-specific screening data were obtained from the Finnish Vammala *H pylori* screen and treat project. Data for 15 to 17-year-olds were used in the analysis of the group aged 15 years (n=988), data for 28 to 32-year-olds in the group aged 30 years (n=832) and data for 45-year-olds in the group aged 45 years (n= 478).

A multivariate sensitivity analysis was performed, where all lifetime probabilities to develop *H pylori*-related disease (gastric cancer, gastric premalignant lesions, duodenal and gastric ulcers, dyspepsia) were simultaneously changed to minimum or maximum values.

A probabilistic sensitivity analysis was performed with a Monte Carlo simulation of 1000 cases. For the simulation, a logistic or uniform probability distribution was created for the ranges of the selected 22 key variables. In each case, a parameter value

was randomly selected from the distribution functions. For other variables, baseline values were used. Thus, the expected cost was calculated with the decision tree for each case, and based on these costs, the mean and standard deviation (SD) was derived for both alternatives (Doubilet et al 1985).

6 RESULTS

6.1 Ultrasound screening in pregnancy (I)

The main effect of ultrasound screening was significantly lower perinatal mortality in the screened compared with the control group. Nineteen fewer babies died, with perinatal mortality being 4.6/1000 vs. 9.0/1000; $p < 0.05$. The difference was mainly due to better early detection of major malformations (detection rate 47%) and subsequent induced abortions. Twelve fetuses with major malformations were found in the screening group, which led to eleven induced abortions. In the control group, none of the women had induced abortions due to ultrasound findings (Saari-Kemppainen et al 1990).

The actual realised gross cost of one screening ultrasound examination, was \$86. The major cost items were lost working time cost (25%), staff (23 %), equipment (19%) and overhead costs (17%). The random sample revealed that the women spent 74 min travelling; 81% were working, more than half of whom used their working time for the ultrasound appointment. In the entire sample the average time spent at the outpatient department was 61 min. The actual time reserved for the appointment was 22 min at the City Hospital and 30 min at the University Hospital.

The cost of screening-induced advanced examinations and procedures was 16% of the total gross unit cost of \$102. By adding the later obtained data of more numerous visits to a health center the unit cost became \$ 114. (Leivo et al unpublished results). In the screening group there were more visits to the maternal health center (12.9 vs 12.5, $p < 0.0001$).

The negative cost, i.e., the savings, of screening was due to less use of antenatal health services. Total outpatient visits were fewer in the screening group (mean 2.34 vs. 2.59; $p < 0.0001$), as were inpatient days (mean 2.32 vs. 2.51, $p = 0.15$). The mean number of other antenatal ultrasound examinations was also smaller, at mean 1.14 (excluding the screening ultrasound) vs. 1.80; $p < 0.0001$. The total negative unit cost calculated per screening ultrasound examination was \$182, two-thirds of which was accounted for by the lower inpatient cost.

The gross costs of avoiding one perinatal death were US\$ 21 938, when the changes in the use of other health services were excluded. The ratio was US\$ 19 096, when the savings from fewer outpatient visits were taken into account. Adding the savings from fewer inpatient stays, the ratio was US\$ -7100, while also including the savings from fewer ultrasounds made it US\$ -17 077. When also was added, the later obtained data of costs of more numerous visits to maternal health centers the ratio was

US\$ -11 515.

The University Hospital used 8 min longer per ultrasound examination and a more detailed pre-structured form for findings than the City Hospital. The malformation detection rate was better in the University than in the City Hospital (75% vs. 35%), and the gross unit cost for ultrasound was \$124 in the former and \$89 in the latter. The cost-effectiveness ratio by hospital was estimated as gross costs, (i.e., excluding the use of other health services) vs. induced abortions. This ratio was \$28 005 in the University and \$55 654 in the City Hospital.

Even substantial changes in gross costs changed the screening cost-effectiveness ratio very little. By contrast, the cost-effectiveness ratio changed +19%, when the savings of antenatal inpatient days was changed by 5%. Changing the values of other savings items had only a marginal effect.

6.2 Breast carcinoma screening (II)

The relative risk of death from breast cancer among the invitees compared to those not invited was 0.76 (95% confidence interval 0.53-1.09), which indicated a 24% protective effect due to screening which was not statistically significant (Hakama et al 1997). The total undiscounted LYS in the sample of 89 893 invitees in 1987-1992 were 56.

In the base case, where the difference in cumulative relative breast cancer mortality between the screening and control group was estimated to remain constant after 5 years following diagnosis, the total number of discounted life-years saved was 578. In 1987-1992 the number of discounted life years saved was 46, i.e., 8.0% of the total life-years saved. In the second case, where the difference in cumulative relative breast cancer mortality between the groups was estimated to be constant after 7 years following diagnosis, the total number of discounted life years saved was 585. In the third case where the difference in cumulative relative breast cancer mortality between the groups was assumed to be constant after the end of the screening in 1992, the total LYS was estimated to be 271.

During 1987-89, 76,389 women entered the program, and on the assumption of a 90% participation rate in the re-screenings of 1990 to 1992, altogether 201,141 screening examinations were performed. The unit health care cost of screening (excluding further assessment) was between US\$25 and US\$45. This difference was mainly due to differences in the salary and social expenses cost. Taking into account the cost of assessment, the unit cost of screening was between US\$28 and US\$51. The screening unit's higher unit cost of screening did not correlate with higher quality of screening, measured by the cancer detection rate (Leivo unpublished result). The unit cost difference reflects the actual resource cost, which differed from the market price of the

service. The high cost screening unit set the market price below the resource use cost, and subsidized the screening by other sources. (Leivo unpublished result).

The unit cost of transportation (one-way) was US\$3 (SD US\$4) and the cost of time was US\$4 (SD US\$12). The average total time spent was 1 hour 12 minutes.

The unit health care cost of surgery for false positives was on average US\$1,548. The cost of transportation was US\$8 and the time cost US\$267, resulting from the normal practice of giving one week's sick-leave.

The health-care saving per every new case of breast cancer due to early treatment was estimated to be US\$2,091; transportation cost was US\$20 and time cost US\$325.

The national ratio of carcinoma in situ ductal to total breast carcinomas was 1.54 percentage units higher and the ratio of carcinoma in situ non-ductal to total breast carcinomas was 0.93 percentage units higher in the period 1990-94 compared to the period 1980-84. The 1990-1994 figures were applied to the screening group. The cost of treating 72% (Page et al 1982) of those new intraductal breast carcinomas and 62% (Rosen et al 1978) of those new nonductal breast carcinoma in situ, which exceeded the proportion of equivalent carcinomas in situ to carcinomas in total estimated from the 1980-1984 data, was assumed to be the cost of the overtreatment of breast carcinoma in screening. Consequently, we estimated 10.1 ($1.54 \% * 907 * 72 \%$) ca in situ ductal and 5.2 ($0.93 \% * 907 * 62 \%$) ca in situ non-ductal to be due to over-diagnosis in the screening group.

The total costs of the screening in the base case were US\$11 million, i.e., US\$14.3 million per 100,000 participants, with 84% of the total cost consisting of the base mammography screening. Altogether 82% of the total costs were health care costs.

The discounted LYS per 1,000 screens was 3.2 in the base case. The cost per life-year saved was US\$18,955 in the base case.

In the univariate sensitivity analysis the C/E ratio was sensitive to changes in the number of life-years saved and to the base cost of the mammography. However, the C/E ratio was not sensitive to changes in other cost items.

Under different model assumptions the cost-effectiveness ranged from US\$15,502 to US\$40,308. The largest deviation caused by the change in costs, +30% compared to the base case, fell under assumption IV, i.e. when time cost was counted for all attendees irrespective of working status. The largest deviation caused by the change in effects, +123% compared to the base case, was under assumption IX, meaning when

the cumulative relative breast cancer mortality reduction was assumed to be constant after the end of screening (1992) in the study and control groups. (Table 6).

Table 6 - Total costs (US\$), effects and cost-effectiveness of breast cancer screening under various model assumptions

	Costs	Effects	C/E
I base case	10 956 115	578	18 955
II surgery cost only for clinically evaluated unnecessary false positives	9 241 836	578	15 989
III no overtreatment cost	10 737 529	578	18 577
IV time cost for all attendees; based on average salary of sample	14 260 234	578	24 672
V no transportation and time cost	8 959 890	578	15 502
VI discount rate 5%	10 302 932	431	23 905
VII no overtreatment, transportation and time cost, discount rate 5%	8 254 462	431	19 152
VIII difference in breast ca survival constant after 8 year from diagn. in screening and no-screening groups	10 923 341	585	18 672
IX difference in breast ca survival constant after end of screening (1992) in screening and no-screening groups	10 923 341	271	40 308

II According to a Finnish study, 30% of these operations should be classified as unnecessary, whereas the rest of the lesions have a small, but elevated risk of malignancy or might cover a malignant lesion (Jahkola et al 1991)

6.3 Double-reading mammograms (III)

The total number of cancers detected with the double and single reading was 290 and was 261. The proportions of cancers detected by the second reader were closely similar in all three screening centers, with no statistically significant difference ($p>0.1$). Differences in possible consensus practice or in number of further assessment procedures had no meaningful effect on the proportion of cancers detected among the mammographs performed.

Cancers detected by the second reader only were slightly smaller in size by mammography, but were larger by microscopy. These differences were not statistically significant. A significantly ($p<0.05$) higher ratio of carcinoma in situ was the causative pathology in cancers detected only by the second reader (40% vs 17%).

The incremental unit base cost of screening was US\$ 6. The labor cost of the second radiologist was US\$ 5. The rest were the allocated costs of the additional office work, space and overhead costs. The incremental total base cost of screening was US\$ 589 570, representing 80% of the incremental total net cost of double reading.

The incremental total cost of double reading was US\$ 740 160, representing 13.0% of the total screening programme cost of US\$ 5.5 million with the double reading strategy.

The incremental assessment procedures cost caused 27.2% of the total assessment-procedures cost, and the incremental surgical procedures cost was 13.2% of total surgical-procedure cost.

The cost per one cancer detected with a single reading was US\$ 18 340 and with a double reading US\$ 19 058. The incremental cost of an additional cancer found was US\$ 25 523, i.e., a 39% higher cost per cancer found by the double compared to the single reading.

The incremental C/E ratio varied from US\$ 21 026 to US\$ 34 744 in the different screening centers. The variation was mainly due to differences in the number of incremental cancers found in each screening center and partly due to differences in the number of assessment procedures performed.

In the univariate sensitivity analysis the C/E ratio was clearly sensitive to the number of cancers detected and to health care cost of the base screening test. It was not sensitive to changes in other cost items. When only invasive cancers were included and carcinoma in situ cases were excluded, the C/E ratio was US\$ 42 538.

6.4 Helicobacter pylori screening (IV)

The incremental health care cost per case under different multivariate model variations is given in Table 7. The cost per case, i.e. person invited for screening, was US\$ 69 in screening and US\$ 43 in the no-screening alternative in baseline case analysis. Thus, the incremental cost per case was US\$ 26. Incremental cost per treated H pylori infection due to screening was US\$ 412. The incremental cost per case was highest in the group aged 15 years and lowest in the group aged 45 years, where the H pylori screening showed cost savings per case. The incremental cost per case was US\$ 38 when all disease probabilities were simultaneously changed to minimum values and US\$ 15 when maximum values were used.

In the probabilistic sensitivity analysis, the screening vs. no-screening values were: mean 68.7 vs. 43.1, SD US\$ 4.5 vs. 5.3, minimum values US\$ 55 vs. US\$ 27 and maximum values US\$ 81 vs. US\$ 63.

In the univariate sensitivity analyses, the largest deviations were caused by changes in the estimated probability of H pylori-caused gastric cancer (cost benefit range from US\$ 24 to US\$ 30) or gastric pre-malignant lesions (range from US\$ 23 to US\$ 31). The cost difference between the screening and no-screening alternatives was not sensitive to changes in the selected cost variables, notwithstanding the baseline cost of the screening test and the GP visit (Table 7).

Table 7. Incremental costs (US\$ 1998/99) per case under selected model variations.

Model	US\$
<u>Multivariate analysis</u>	
Baseline case	26
Age at screening 15 years	36
Age at screening 30 years	27
Age at screening 45 years	-6
All Hp-related disease probabilities at minimum value	38
All Hp-related disease probabilities at maximum value	15
Base cost of screening +20% (Hp test + Gpvisit)	31
<u>Univariate analysis</u>	
Estimated lifetime probability of Hp-related cancer 0.005, if Hp+	30
Estimated lifetime probability of Hp-related premalign 0.020, if Hp+	31
Screening test specificity 0.98	20

Hp = Helicobacter pylori

VII DISCUSSION

7.1. Material and methods

7.1.1 Cost-effectiveness-analysis material and methods (I, II, III)

In assessing screening programmes, process measures instead of direct outcome measures are often chosen. The choice of process measures causes more uncertainty to the economic evaluation of screening than the choice of final health output measures. In the present study the measure of cost-effectiveness was related to the final health output, i.e. life-years saved in the breast carcinoma screening study (II). In the double-reading mammograms study (III) the measure of cost-effectiveness was incremental cost per additional cancer found. An additional case of cancer found is only an indirect measure of the life years saved. Due to missing data, the number of life-years saved could not serve as a measure of effectiveness. No published clinical trial study could be detected, where the number of life-years saved due to double -reading would have been estimated.

It is noteworthy that the one-stage ultrasound screening (I) does not inevitably produce life-years saved, although it reduces perinatal mortality. It can even produce less life-years. The cost of avoiding one perinatal death gives us the cost of the mainly psychological benefit of the mother being able to terminate the pregnancy instead of giving birth to a dead baby, a baby who would die immediately after the delivery or who would be severely malformed.

Cost-effectiveness of the ultrasound screening (I) was strongly dependent on the rates of malformation detection and subsequent pregnancy termination. It was implicitly assumed that the perinatal mortality rates in the control and intervention groups would have been the same if no screening had taken place. Although the patients were randomly assigned to the groups, this assumption may not be justified. In the screening group 19 fewer babies died, with perinatal mortality being 4.6/1000 vs. 9.0/1000; $p < 0.05$. However, the number of major malformations found was 12 and the number of ultrasound induced abortions was only 11. In this aspect, the choice of cost per major malformation found would have been more unambiguous measure of cost-effectiveness.

The ultrasound study (I) effectiveness data were considered representative being from a randomized, controlled trial, which included a large sample, lost only four women

during follow-up, and had complete mortality data, all of which led to accuracy in the estimation of effectiveness.

The breast carcinoma study (II) effectiveness data were from a nationwide population-based screening programme from 1987 to 1992 and from an estimation model using parameters based on published studies and national cancer statistics from 1993 to 2020. Finland was the first country to introduce nationwide breast cancer screening as a public health policy; this was in 1987. The program was expanded rapidly, and thus the follow-up of the effect could be evaluated only to the end of 1992, because the controls were also gradually being screened (Hakama et al 1997). Consequently, only 8% of the total LYS emerged during the actual follow-up period in our base-case estimation. That, on average, the follow-up was four years and there was a statistically non-significant effect, reduced markedly the possibility of estimating the exact number of life-years saved. The effectiveness data were subject to considerable limitations due to the time horizon. As a consequence, the economic evaluation was drawn on published articles and methods of extrapolation, i.e. modeling.

The effectiveness data of double-reading mammograms study (III) were original retrospective data. They were considered representative, being community-based and comprising 15% of all mammographs read in Finland during a 5-year period within a nationwide screening program. We had access to the original gradings of the first and the second reader, which enabled us to estimate the incremental effectiveness of the double-reading accurately. The fact that the second reader had access to the results of the first reader's results means that the incremental nature of the evaluation does not fit the definition of independence. However, our study evaluates the incremental cost-effectiveness of the actually used double reading strategy in Finland, i.e., a strategy where the double reading was not all blind.

The costs were evaluated from the societal perspective, i.e. including the health care and non-health care and time costs of the screening intervention. In the present studies the non-health care cost was the cost of transportation. Time costs did not include the time cost of possible accompanying persons at the screening. This cost was estimated as insignificant due to the age of screenees. The cost of lost home work was only included in part of the sensitivity analyses, but in no base line cases. The valuation principles of the opportunity cost of unpaid inputs are under debate among health economists (Posnett et al 1996) and we chose to exclude the cost of lost home work in base line cases, since that is the most common practice in current published cost-effectiveness analyses. However, this practice does not follow the principle of broad societal perspective. Additionally, following the guidelines by Gold et al (1996) and common scientific practice, productivity costs were not included in the cost accounting, although contrary opinions exist among health economists (Brouwer 1997). The valuation of extremely intangible and ambiguous costs of screening, as the cost of anxiety, reassurance and information per se were not either included in the analyses. The inclusion of cost of lost home work, productivity cost and costs of anxiety would have increased the present cost-effectiveness ratios. The inclusion of the costs of reassurance and information per se would have decreased these ratios.

The cost savings due to terminating a life were not included in the cost-effectiveness analysis of ultrasound screening in pregnancy. This was due to the short time horizon of the analysis. Although morally difficult, a longer time horizon and subsequent inclusion of the saved resources associated with the termination of pregnancy would theoretically have produced more detailed estimates for the economic evaluation.

Actual resource use costs were used, when evaluating the base cost of screenings and most of the screening induced further examinations. This method sharpens the cost estimation significantly, as Finland is a Nordic welfare state, where the health care is mainly publicly financed and the competition among producers is scarce and often distorted. The pricing of products or services is seldomly resource-based. The actual resource use costs were obtained from the screening organizations' internal accounting. Various adjustments were made in order to standardize the different accounting methods. This was done in co-operation with the organizations' accounting managers in order to obtain all the data behind the public figures and thus to produce the most reliable resource use estimates.

The antenatal health care (I) is traditionally not based on competition between different producers. Thus we used the internal accounting data from the two trial public hospitals, where the screening took place. At the time of the cost estimation of ultrasound screening the accounting principles in the two hospitals did not comprehensively include many major cost items, e.g. the rent, overhead, depreciation or capital costs. These were adjusted for in our study estimates. Compared to most Finnish health care interventions the mammography screening (II, III) has been a competition between different producers at least in some municipalities. However, the quality of the cost accounting methods in the different organizations was variable, and thus adjustments were made to standardize the different accounting methods. This sharpened the cost estimation significantly.

Our studies did not include the quality adjusting of life-years gained. In the ultrasound study the measure of cost-effectiveness was cost per avoided perinatal death, which as such can not be quality-adjusted. The quality-adjusted life-years saved of mammography screening need to be evaluated in future studies.

7.1.2 Cost-benefit – analysis material and methods (IV)

Cost – benefit analysis using computer-based decision tree was chosen for the method for economic evaluation of H pylori screening, as for the time being no randomized controlled trial data, or other clinical data exist for the evaluation of life-years saved due to H pylori screening. Due to the time horizon, the available Vammala Helicobacter pylori trial data was not suitable for the evaluation of effects measured in life-years saved. In addition, the study was conducted to answer a specific question by

the clinicians and decision makers concerning the possible health care cost savings due to nationwide H pylori screening programme. Few primary studies will be able to collect direct data measuring effects along the entire cascade of events from intervention to health outcome. Modeling must often be used to combine information from different data sources. (Gold et al 1996).

Cost – benefit analysis measures costs and benefits in monetary terms. In theory this implies that all the consequences of an intervention would be expressed in monetary terms. In practice, cost-benefit analyses usually amount to a comparison of those costs and benefits that can be easily expressed in money terms (Drummond et al 1987). In the present study productivity costs and the intangible costs of a more diffuse nature, such as the value of death or the value of anxiety, reassurance or information per se were excluded from the analysis. Thus the willingness to pay of the individuals who benefit from the programme was not estimated. This method is recommended by e.g. Johannesson et al. (1996), but in the present study our aim was to include only unambiguous monetary costs in the analysis.

The non-health care and time costs and benefits were not included in the present model. Their inclusion would have increased the uncertainty of the outcome of the model. Unfortunately, their exclusion makes the comparison between the *Helicobacter pylori* – study (IV) and the other studies (I, II, III) more difficult. This cost-benefit analysis is a first-line approach taking into account only the direct financial implications of the H pylori screening. However, screening also produces better health. The cost-effectiveness, life-years saved and quality-adjusted life years saved from broad societal perspective need to be evaluated in future studies.

The decision tree model was planned keeping in mind the complexity of the problem and the clarity of the model. For the sake of clarity, the H pylori - related disease probabilities were presented as net probabilities. The probability of a H pylori positive patient to develop a H pylori-related disease due to infection was presented instead of distinct probabilities of H pylori positive and negative patients to develop the disease. Certain cost-causing events, e.g., costs prior to the H pylori – related disease diagnosis, were not represented as nodes in the decision tree, but included directly in the aggregated cost estimates. The estimated lifetime risks of H pylori-related diseases were disease-specific in the model, i.e., the possibility of successive H pylori-related diagnoses was not included in the model. Thus the probability of no clinical symptoms may be underestimated. Possible re-infection or spontaneous cure were also excluded. All these simplifications undoubtedly impact on the outcome of the model.

In the decision model structure after participation the following branch was infected / not infected, and the following branches sensitive/ not sensitive test and specific/ not specific test. This may cause slight incorrectness of the outcome of the model. A more accurate option would possibly have been to plan the model structure by using branch positive test / not positive test and the following branches true positive/ false positive and true negative / false negative.

Probability estimates for screening – related variables were considered representative being based on realised pilot screening programme in Finland. However, probabilities may vary markedly across countries (Stone et al 1998). The sensitivity and specificity of the *Helicobacter pylori* screening test were obtained from the Vammala *Helicobacter pylori* trial data. Unfortunately, the confidence intervals of sensitivities and specificities were not represented. However, these both were included in the sensitivity analyses to estimate the impact of possible uncertainty caused by them. The simplifying assumptions of no third treatment failure and 100% sensitivity and specificity of the screening retests were estimated to have little impact on the outcome due to relatively small number of cases.

Probability estimates for *H pylori*-related disease variables were obtained from selected references in the literature. Because some of the *H pylori*-related conditions and consequences are little studied, all probabilities could not be based on solid existing evidence thus causing possible uncertainty in the baseline outcome of the model. Consequently, in some aspects the probability estimates were based on expert opinion, which can be affected by various biases. (Poses et al 1988). Attempts to compensate for this, include carrying out several sensitivity analyses and using relatively wide variable ranges in these analyses.

The evidence for potential negative effects of curing *Helicobacter pylori* is controversial. We have not considered the potential adverse effects of wide use of antibiotic therapy leading to bacterial resistance. In addition the model does not consider the possible effect of the cure of *H pylori* infection on the development of reflux esophagitis and the primary adenocarcinoma of the esophagus (Labenz et al 1997). These choices may cause underestimation of the outcome of the model.

7.1.3 Valuation and discounting methods (I, II, III, IV)

The year of price level varied from 1990 (I) to 1995 (II, III) to 1998/99 (IV) in the present studies. It is noteworthy that the early years' cost- effectiveness ratios would be higher at the later years' price levels.

There is broad agreement that in cost-effectiveness analyses all future costs and health consequences should be stated in terms of their present value, i.e., discounted. Only then will the interventions' cost-effectiveness ratios be appropriately adjusted for the decision-maker's time preference for present over future outcomes (Gold et al 1996, Drummond et al 1987). A discount rate of 3% for costs and effects was used in the base case analyses, because given currently available economic data, 3% is recommended as the most appropriate real riskless discount rate for economic analyses by the Panel on Cost-effectiveness in Health and Medicine appointed by the U.S. Public Health Service in 1993 (Gold et al 1996). A discount rate of 5 % for costs and effects would have increased the cost-effectiveness and the cost-benefit of

screenings. A smaller discount rate for the effects than costs would have shown the cost-effectiveness to be more favourable to screening than the current analyses.

7.2 Ultrasound screening in pregnancy (I)

The earlier economic evaluations of screening ultrasound (Ewigman et al 1993, Persson et al 1983, Temmermann et al 1991, Bakketeig et al 1984), which reported the unit cost of screening ultrasound being between \$23 and \$540 (the latter including two ultrasounds) are difficult to compare with our study, because the studies did not estimate the effectiveness. In addition, in the Stockholm trial (Waldenström et al 1988) patients with an indication for ultrasound were excluded, and the Norwegian trials (Bakketeig et al 1984) had small population sizes. It is worth noting, that these studies did not report the cost-accounting principles, which makes also the comparison of the unit cost difficult. The use of other antenatal health services has been very variable in other studies from more use to less use (Persson et al 1983, Jørgensen 1992, Backe et al 1990). The use of other antenatal health services is most probably different in different health care settings, and as such must be nation-specifically estimated.

DeVore (1994) reported the cost per detected malformed fetus as \$29 533 in tertiary centers and \$115 575 in non-tertiary centers, depending on the different diagnostic rate of malformed fetuses in different types of centers. Similar to the present study the cost-effectiveness ratio was better in the centers, where the detection rate for anomalies was highest. It is noteworthy that in the DeVore (1994) study the measure of cost-effectiveness was the cost per detected malformed fetus, which decreases the C/E ratio compared to our study. DeVore (1994) used the effectiveness data from RADIUS study. In the RADIUS trial (Ewigman et al 1993) around 60% of all those registered were excluded, based on various risk-factors, which can markedly change the outcome of the trial, whereas the Helsinki trial was population-based. In the Helsinki ultrasound Trial the detection rate of major malformations was 40% in the screened group overall. The RADIUS study had a detection rate of 16.6% before the 24th week, the total detection rate being 35%. The different C/E ratios reflect also the different methods of organisation, i.e., the RADIUS - study was done in the United States, the ultrasound screening was done twice and it was performed by obstetricians. These differences increase the C/E ratio compared to our study.

It is of note, that even though the present analysis is based on a randomized controlled trial, the decrease in perinatal mortality was substantially higher than the number of ultrasound induced abortions (19 vs. 11) in the screening group. This may cause uncertainty about the effectiveness, which is not totally indicated in the present study.

I consider the gross cost-effectiveness ratio (US\$ 21 938 per one avoided perinatal death) , i.e., excluding the changes in the use of other health services as the most unambiguous outcome of the present study. However, the screening was cost-saving,

when all antenatal use of other health services were included. It is noteworthy that the cost-effectiveness of the ultrasound screening was strongly dependent on the negative cost of inpatient days. However, the difference in inpatient days between the groups was not statistically significant. Any difference between the inpatient days of the two groups is possibly therefore a result of chance, and the cost-effectiveness findings concerning inpatient days uncertain.

7.3 Breast carcinoma screening (II)

Comparison of our results with other breast cancer C/E studies is difficult, mainly due to different methodological principles, but also due to national differences concerning breast cancer incidence, survival and health care organisation. The year of estimation also has an effect on the cost per life-year saved due to inflation. Previously published C/E ratios are given in Table 1.

Studies which have reported unclearly the LYS (Gerard et al 1990, Forsyth 1988) or have estimated the LYS for a relatively short period after screening are difficult to compare with our study (Australian Health Minister's Advisory Council 1990, Clarke et al 1991, Forrest 1987, U.S. Congress, Office of Technology Assessment 1987, Kattlove et al 1995, Hristova et al 1996). Gerard et al (1990), Forsyth (1988), Okubo et al (1991), Kattlove et al (1995), and Hristova et al (1996) did not discount the LYS, which led to considerable underestimations of the cost of LYS.

The majority of previous studies have used a 5% discount rate compared to our 3% (Australian Health Minister's Advisory Council 1990, Clarke et al 1991, Forrest 1987, U.S. Congress, Office of Technology Assessment 1987, Eddy 1989, Brown 1992, van der Maas et al 1989, de Koning et al 1991, Hall et al 1992, Lindfors et al 1995, van Ineveld et al 1993). To facilitate the comparisons we estimated the cost-effectiveness of the base case also with a 5% discount rate, which resulted in a 26% higher cost-effectiveness ratio compared to the base case. Thus, a base case discount rate of 5% in the present study would have increased the cost per life-year saved markedly.

Most authors have estimated the breast cancer mortality reduction to be over 28% (Forrest 1987, U.S. Congress, Office of Technology Assessment 1987, van der Maas et al 1989, de Koning et al 1991, Lindfors et al 1995, van Ineveld et al 1993, Hristova et al 1996, Norum 1999). Due to this our estimation of the cost per life year saved is higher than in other studies.

The cost-effectiveness of breast cancer screening is very sensitive to the LYS, and thus to the methods used in estimating the LYS. In 1987, Finland became the first country to introduce nationwide breast carcinoma screening as a public health policy. The program was expanded rapidly, and thus the follow-up of the effect could be evaluated only to the end of 1992, because the controls were also gradually being screened. Consequently, only 8 % of the total LYS emerged during the actual follow-

up period in our base-case estimation. The fact that, on average the follow-up was 4 years and was a statistically non-significant effect, reduced markedly the possibility of estimating the exact number of LYS. According to most studies based on clinical trials or case-control studies, a major proportion of the deaths prevented will be seen after the first 4 years. However, the size and length of the effect, as well as the screening schemes, are variable. The LYS depend crucially on whether early or late deaths are prevented, and in this regard most published trials have relatively short follow-up times.

The majority of published trials and meta-analyses have shown screening mammography to reduce breast carcinoma mortality among women older than 50 years. However, the Canadian, Edinburgh and Finnish trials report either no effect or a statistically insignificant effect (Miller et al 1992, Kerlikowske et al 1995, Hakama et al 1997, Roberts et al 1990, Alexander 1997). Meta-analyses by Gotzsche et al (2000) reported mammography screening not to reduce breast cancer mortality. If the truth about mammography lies close to Gotzsche, the cost/LYS would be much higher than in the present study base case, or even the present study maximum case, derived from the sensitivity analysis. In the extreme case, where effectiveness would be zero, the cost – effectiveness would be infinite.

The cost-effectiveness of breast cancer screening is also sensitive to the base cost of screening. The base cost of screening in Finland (US\$ 42), even though contrary to most published studies (U.S. Congress, Office of Technology Assessment 1987, Eddy 1989, Brown 1992, Okubo et al 1991, Kattlove et al 1995, Hall et al 1992, Lindfors et al 1995, van Ineveld et al 1993, Hristova et al 1996) the non-health care and time costs were included, was only average.

The C/E ratio in our study estimated the cost-effectiveness of the national program during the period 1987-1992. The C/E ratios, including different age-groups, screening intervals, and quality-adjusted life-years, need to be evaluated in future studies. The quality-adjustment would increase the cost per LYS, in the previous studies the multiple has been 1.01-2.27 (Australian Health Minister's Advisory Council 1990, Clarke et al 1991, Forrest 1987, de Koning et al 1991, Hall et al 1992, Hristova et al 1996).

Methodologically our study is quite comparable with the studies of the U.S. Congress, Office of Technology Assessment (1987), de Koning et al (1991) and van Ineveld et al (1993). In the OTA study the screening was annual, and both screening costs and effects were evaluated for the years 1990 to 2020, after which no costs or effects were estimated. The discount rate was 5%. These assumptions increase the cost per life-year saved compared to our base case estimation. De Koning et al (1991) and van Ineveld et al (1993) have used the same computer model when estimating cost-effectiveness. Van Ineveld et al (1993) has simulated the cost-effectiveness in the United Kingdom, the Netherlands, France and Spain. In Spain the incidence and survival rates are relatively similar to those in Finland, and the estimated discounted

life-years saved per 1,000 screens (3.9 vs.3.2) and cost per life year saved (US\$15,326 vs. US\$18,955) were similar.

The inaccuracy in the estimation of life-years saved is mainly due to the fact that the cost of breast carcinoma mammographic screening per life-years saved varied under different model assumptions in the present study. Taking into account the range from US\$ 15 502 to US\$ 40 308, it can be concluded that the base case estimation of US\$ 18 955 represents more the low than high estimate of the cost - effectiveness ratio.

7.4 Double-reading mammograms (III)

In the present study a double reading was more effective in breast cancer detection than a single reading. These findings confirmed those reported elsewhere (Anttinen et al 1993, Anderson et al 1994, Brown et al 1996, Ciatto et al 1995, Warren et al 1995, Brown et al 1996, Blanks et al 1998).

Brown et al (1996) found saving per additional breast cancer detected of GBP 485 (US\$ 766) when comparing consensus double with non-consensus single readings, and the incremental cost per additional cancer detected was GBP 1926 (US\$ 3043) when comparing non-consensus double to non-consensus single reading. They estimated that the incremental unit health care cost of the second reading would range from GBP 0.62 to 0.69 (US\$ 0.98 to 1.09), based on the observation that a total of 1 980 images were read in 678 minutes. This exceeds the Finnish reading speed markedly. In addition, we have included in our calculations the total working time paid for, not just the effective performance time. We also included the saving due to the early detection included in the majority of the breast cancer cost-effectiveness studies. Unlike Brown et al (1996), we focused solely on the cost-effectiveness of the double compared to the single reading, and assumed a strategy of single reading with consensus. A single reading strategy does not inevitably exclude consensus reading.

The biological significance of carcinoma in situ in breast cancers detected by screening is uncertain. A higher incidence of carcinoma in situ in the double-reading than in the single reading screening strategy may be a sign of over-diagnosis or earliness of diagnosis (Wright et al 1995, Ernster et al 1996). If the carcinoma in situ lesions were due to over-diagnosis, than there was in fact a higher incremental cost for additional cancer found by the double reading strategy. If all the additional carcinoma in situ are considered over-diagnosis, the base case incremental C/E ratio would be nearly twice as high (25 523 US\$).

7.5 *Helicobacter pylori* screening (IV)

Comparison of the present study with previous studies is limited due to two main factors. Firstly, no previous study has represented the economic evaluation of *H pylori* screening taking into account all the main *H pylori*-related diseases. Secondly, the previous economic analyses have been carried out as cost-effectiveness instead of cost-benefit analysis.

Parsonnet et al (1996) and Fendrick et al (1999) have estimated the cost-effectiveness of *H pylori* screening taking into account one *H pylori*-related disease, gastric cancer. Compared to the present study the main difference in the cost accounting methods of Parsonnet et al (1996) was, that they included the future resource use costs in added years of life, and of Fendrick et al (1999) was, that they did not include the treatment cost of gastric cancer in their baseline analyses. Both of these increase the cost-effectiveness of screening. There is a longstanding theoretical controversy about including future resource use costs for unrelated illness in added years of life. In addition, there are many practical difficulties in including these costs (Gold et al 1996).

However, our results, like those of Parsonnet et al (1996) found that cost – effectiveness or cost- benefit was less favorable, if screening was performed at younger ages.

It is noteworthy, that to date there is not much data available from definitive studies quantifying the risk of *H pylori*-related diseases. Consequently, all probabilities could not be based on solid existing evidence thus causing possible uncertainty about the outcome of the model. Attempts to compensate for this included carrying out several sensitivity analyses and using relatively wide ranges in these analyses. However, the choice of variable range is nevertheless subject to discussion as there is little direct evidence that eradication of *H pylori* will reduce the risk of Hp-related diseases.

VIII SUMMARY AND CONCLUSIONS

The purpose of the present study was to conduct an economic evaluation of two nationwide Finnish screening programmes, and one pilot Finnish screening programme intended for nationwide use.

Cost –effectiveness analysis was used for the evaluation of ultrasound screening in pregnancy, breast carcinoma screening and double-reading mammograms practice within breast carcinoma screening. The outcome measures were the cost per avoided perinatal death due to ultrasound screening, cost per life-year saved due to breast cancer screening and incremental cost per additional cancer found due to double reading mammograms. Cost-benefit analysis was used for the evaluation of *Helicobacter pylori* screening using a computer-based decision tree. The outcome measure was the incremental cost per screenee.

The effectiveness data of the ultrasound screening came from the Helsinki randomised ultrasound trial comprising a trial population of 9310 pregnant women. The mammography screening effectiveness data came from Finnish nationwide population-based screening programme, comprising 90 000 invitees and 69 000 controls. The data was completed with an estimation model. The double-reading mammograms effectiveness data were originally obtained from three Finnish screening centers, comprising 95 000 mammograms performed during 1990-1995. Probability estimates for screening-related variables were obtained from the Vammala *Helicobacter pylori* pilot screening study comprising 5300 screenees. Probability estimates for *Helicobacter pylori* related disease variables were obtained from published reports and national statistics.

Health care, non-health care and time costs were included in the cost-effectiveness analyses, only health care costs were included in the cost-benefit analysis. Cost-accounting data were obtained from internal accounts of screening units, published studies, national statistics, health market sources and a questionnaire completed by a sample of 500 women at the ultrasound screening and by a sample of 1400 women at the mammography screening.

The discount rate was 3%. Several univariate and multivariate sensitivity analysis were performed to reveal the robustness of the results to changes in key variables.

The cost per avoided perinatal death was US\$ 21 938, when changes in the antenatal use of other health services were excluded. The screening became cost-saving when antenatal use of other health services was included in the calculations. It is noteworthy that the cost-effectiveness of the ultrasound screening was strongly dependent on the savings from fewer inpatient days. The difference in inpatient days between the

groups was not statistically significant and the cost-effectiveness findings concerning inpatient days uncertain. However, the present study suggests that longer ultrasound examination time and more numerous advanced examinations are rewarded by a better cost-effectiveness ratio. This indicates that high quality screening is of extreme importance, not only from ethical, but also from a cost-effectiveness perspective. The second trimester ultrasound screening during pregnancy can be considered justifiable. It is of note that the information per se which was not accounted for in the present study is also of major utility to the pregnant women. Special care has to be taken, that the women are well informed about the screening, and that the principle of self-determination concerning screening and possible screening induced abortions is strictly respected.

In breast cancer mammography screening the cost per life-year saved was US\$18 955 in the base case, ranging from US\$15 502 to US\$ 40 308 per life-year saved under different model assumptions. The inaccuracy in the estimation of life-years saved was mainly due to the variation of the outcome under different model assumptions. The base line case should be regarded as being towards the lower rather than higher end of estimates of the cost-effectiveness of breast cancer screening. The cost-effectiveness is strongly dependent on the estimated effectiveness, i.e. number of life-years saved. A major controversy still exists in the scientific community over the amount of effectiveness of mammography screening in the prevention of mortality. Some researchers have reported mammography screening not to be effective. If this is the case, the possible range of cost/ life-year saved is much wider than in the present study. To conclude, mammography screening is a high-cost intervention with to some extent uncertain effectiveness. Nevertheless, in the female population, breast cancer is the most common cancer, with substantial mortality irrespective of modern treatments, making decision making about mammography screening morally very difficult.

In double-reading mammograms the cost per additional cancer found was US\$ 26 000. The incremental cost per cancer found was 39% higher with double reading compared to single reading. Double-reading found 10% of all the cancers. A significantly higher ratio of carcinoma in situ was the causative pathology in cancers detected only by the second reader. These results suggest that double-reading mammograms is more effective, but also markedly costlier than single reading. Decision-maker could consider allocating the resources to other use within breast cancer screening intervention or other interventions.

The H pylori screening resulted in an extra cost of US\$ 19 per screenee in the base case of the age group 15-45 years old. The incremental cost per screenee was highest in the 15 years' age group and lowest in the 45 years' age group. These results are subject to considerable uncertainty due to the lack of published clinical data. The present results suggest that the *Helicobacter pylori* screening is not a cost saving intervention, but that the incremental cost of H pylori screening may be surprisingly low. Further pathophysiological and evaluative studies in the area are encouraged. In

the 15-45-years-old population the possible screening should be targeted at the older cohorts.

The use of economic evaluation when making decisions about the broad allocation of resources requires comparisons of a wide range of interventions. At present only a few evaluations of screening, and respectively of other health interventions, are available. Antenatal screening and screening in child counseling clinic comprises a major part of the screening as a whole in Finland. It is noteworthy that very few of these screening procedures are evidence-based or have been subject to economic evaluation. Additionally, the screening in child counseling clinic comprises of many interventions, which are not formally structured or verified. In this sector the need for screening for social and psychological disorders is growing (Appendix 1,2). Screening for these kinds of disorders is possible, but new intervention strategies with structured process forms suitable for first-line and continuous evaluation are needed.

Economic evaluation of screening intervention should be a continuous process alongside with the intervention. One of the key difficulties in the present studies was, that the economic evaluation was not considered as a part of the screening process, but the need for it became evident to the decision-makers only many years after the beginning of trials.

For the time being, cost accounting data and methods in the Finnish health care sector are of variable quality. Many researchers have emphasized the fact that the health care cost estimates based on prices are currently mainly not resource-based. Thus, without adjustments they do not serve as a reliable basis for either evaluative studies or competitive markets. It is of utmost importance to develop the cost accounting methods in the public health care in Finland.

In order to make comparisons, standardization of the evaluative methods is essential. To facilitate the allocation of scarce resources, the main principles concerning e.g., the rate of discounting, the methods for valuating health care costs in the absence of resource use cost and the methods of valuating time should be agreed upon. In the future research resources should be allocated for the development of Finnish standards for the conduct of economic evaluations.

APPENDIX 1. CURRENT SCREENING PROGRAMMES IN FINLAND

1. ANTENATAL SCREENING TEST	DISEASE
blood pressure	pregnancy-induced hypertonia pre-eclampsia
urine albumin	pre-eclampsia urine tract infection
urine glucose	gestational diabetes
haemoglobin	anemia of the mother
blood-group antibodies	anemia of the fetus
serum sample for developmental abnormalities screening (b-HCG, PAPP-A, AFP)	21-trisomia meningomyelocele congenital nephrosis
serum sample for infection screening	syphilis hepatitis B HIV
ultrasound 13.-14. or 16.-18. gestational week	structural malformations 21- trisomia twin pregnancies dating of the pregnancy
placental biopsy or amniocentesis (offered only for over 39-years-old women)	chromosomal abnormalities
clinical investigation by midwife (palpation, measuring of the uterus)	fetus growth abnormalities abnormalities of the amount of amnion fluid abnormalities of the fetus position

2. PERINATAL SCREENING TEST

DISEASE

placental serum thyrotropin	hypothyreosis
clinical examination by midwife and pediatrician	see below, 3.

3. SCREENING TEST IN CHILD

DISEASE

GUIDANCE CLINIC

<p>clinical examinations by nurse and GP</p> <ul style="list-style-type: none"> • sight, lamp test • hearing, otoscopic examination • growth • weight • circumference of the head • physical neurological tests • psychomotor developmental tests • cognitive developmental tests • physical testes examination • physical hip examination • physical cardiac examination, including blood pressure 	<ul style="list-style-type: none"> • vision disabilities, squinting • hearing disabilities, chronic middle ear infections • general pediatric diseases • general pediatric diseases • hydrocephalus • neurological diseases and developmental disorders • neurological diseases, psychomotor developmental disorders • neurological diseases, cognitive developmental disorders • undescended testes, retractile testes • congenital dislocation of the hip • congenital heart disease
--	---

4. SCREENING TEST IN DEFENSIVE DISEASE ARMED FORCES

formulated question sheet	mental diseases (symptoms of) all somatic diseases
height and weight	overweight
vision	vision disabilities
hearing, including otoscopic examination	hearing disabilities, middle ear perforations
physical cardiac examination, including blood pressure	heart diseases
physical stomach and gastroinguinal examination	hernias
physical skin examination	pigment lesions susceptible to chafing
physical leg examination	structural deformities
physical back examination	structural deformities
testes	not defined

5. CANCER SCREENING TEST DISEASE

cervical pap-smear	cervical cancer
mammography	breast cancer

References: Hiilesmaa et al 2000, Viisainen 1999, Leinikki et al 2000, Si mell et al 1997, Yleislääkärin käsikirja 1994, Lääkärintarkastusohje 1997, Teppo et al 2000.

APPENDIX 2. PROPOSED SCREENING PROGRAMMES IN FINLAND

SCREENING TYPE	PROPOSED SCREENING TEST	DISEASE	PILOT PHASE
Prenatal screening	<ul style="list-style-type: none"> • serum-toxoplasma • serum-CMV 	toxoplasmosis cytomegalovirus infection	x
Child counseling clinic screening	<ul style="list-style-type: none"> • not defined • electronic hearing test 	psychological or behavioural disorders severely impaired hearing disability	
Cancer screening	<ul style="list-style-type: none"> • papillomavirustest • serum-PSA • feces occult blood • FSA (foetal sulphoglyco-protein antigen) • serum-pepsinogen I • lung computer tomography 	targeted cervical cancer screening prostate cancer colorectal cancer stomach cancer stomach cancer lung cancer	x x
Infection screening	<ul style="list-style-type: none"> • urine-chlamydia • serum helicobacter antibodies 	veneral chlamydia infection helicobacter pylori infection	x
Targeted screening <ul style="list-style-type: none"> • narcomanes • BRCA1/2 	<ul style="list-style-type: none"> • HIV, hepatitis A,B,C- antibodies • mammography 	HIV, hepatitis A, B, C breast cancer	 x

genedisorder <ul style="list-style-type: none"> • FAP-HNPCC-families • asbestos-patients 	<ul style="list-style-type: none"> • colonoscopy • lung CT tomography 	colorectal cancer lung cancer	x
Genetic screening	<ul style="list-style-type: none"> • serum-AGU-gene- disorder • serum-INCL-gene-disorder • serum-fragile X-gene-disorder 	aspartylglucosaminuria infantile neuronal seroidlipofusinos fragile X - syndrome	x

References: Hiilesmaa et al 2000, Leinikki et al 2000, Teppo et al 2000, Saarelma et al 2000, Robinson 1998.

IX ACKNOWLEDGEMENTS

This study was carried out at the Department of Public Health, Medical Faculty, University of Helsinki, during the years 1993 to 2001. I am sincerely grateful to my supervisors, Professor Harri Sintonen, Ph.D., and Docent Risto Tuominen, D.D.S., Ph.D., for their friendly support and guidance during this study.

I thank Professor Olli P Heinonen, M.D., D.Sc., the former head of the Department of Public Health for his advice in the beginning of the study.

I sincerely acknowledge Associate Professor Ivar Kristianssen, M.D., Ph.D., and Docent Pekka Rissanen, D.Sc., for their thorough and constructive revision of the manuscript.

I thank Tiina Salminen, D.Sc., Anniina Salomaa, M.D., and Aulikki Saari-Kemppainen, M.D., Ph.D., for collaboration and friendship.

I thank Docent Eero Pukkala, Ph.D., for his positive attitude and for his excellent advice on cancer epidemiology.

I thank Associate Professor Timo Kosunen, M.D., Ph.D., Docent Martti Färkkilä M.D., Ph.D., Miika Linna D.Tech., Professor Matti Hakama, D.Sc., Kalevi Auerma, M.D., Kaarina Partanen M.D., Ph.D., and Urpo Saari, M.D., for collaboration.

I thank Tuomo Ylinen, Cand.Med., Stud. Tech., for excellent technical assistance on decision tree modeling.

I thank the radiology nursing personnel in Helsingin Lääkärikeskus, Vantaan Lääkärikeskus, Suomen Terveystutkimus Oy and Screening Units of the Cancer Society of Finland for their friendly assistance in collecting the questionnaires for the attendants at the mammography screening.

I thank Financial Manager Juhani Bruun, Helsinki University Central Hospital, Managing Director Marja-Liisa Kotisaari, Pirkanmaan Syöpäyhdistys ry, Manager Raimo Eskelinen, Suomen Terveystutkimus Oy, Financial Manager Tarja-Liisa Ala-Ilkka, Vantaan Lääkärikeskus Oy, Managing Director Kaj Öhman, Vantaan Lääkärikeskus Oy, Financial Analyst Anneli Jääskeläinen, Helsingin Kaupungin Terveysvirasto and Special Analyst Pirjo Lehtonen, Helsingin Kaupungin

Terveysvirasto for providing me unpublished, internal cost data, which has been of considerable value during this study.

I thank Professor Leila Laatikainen, M.D., Ph.D., for giving me time off from my clinical work at the Department of Ophthalmology, Helsinki University Central Hospital, to perform this study.

I thank my mother Pirkko (1933-1996) and father Veikko for their support and encouraging attitude towards education. I express my warmest gratitude to my brothers Tomi, Timo and Ilmo for their positive support and encouragement. Finally, I express my warmest gratitude for my dear Janne for always believing in me.

Financial support was provided by FinOHTA (Finnish Office for Health Technology Assessment) and the Finnish Cultural Foundation.

Espoo, June 2001

Tiina Leivo

X REFERENCES

- Alexander FE, Anderson TJ, Brown HK, Forrest AP, Hepburn W, Kirkpatrick AE, Muir BB, Prescott RJ, Smith A. 14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening. *Lancet* 1999;353(9168):1903-8.
- Anderson E, Muir B, Walsh J, Kirkpatrick A. The efficacy of double reading mammograms in breast screening. *Clin Radiol* 1994;49(4):248-51.
- Anttinen I, Pamilo M, Soiva M, Roiha M. Double reading of mammography screening films – one radiologist or two? *Clin Radiol* 1993;48(6):414-21.
- Australian Health Minister's Advisory Council. Breast Cancer Screening Evaluation Committee: Breast cancer screening in Australia: future directions. Australian Institute of Health: prevention Program Evaluation Series No 1. Canberra: Australian Government Publishing Service, 1990.
- Axon A, Forman D. *Helicobacter gastroduodenitis*: a serious infectious disease. *BMJ* 1997;314:1430-1.
- Backe B, Nafstad P, Saetnan AR. Reduced use of diagnostic obstetric ultrasound in Norway: result of consensus panel recommending routine screening in pregnancy? *Acta Obstet Gynecol Scand* 1990;69:649-50.
- Bakketeig LS, Eik-Nes SH, Jacobsen G, Ulstein MK, Brodtkorb CJ, Balstad P, Eriksen BC, Jorgensen NP. Randomized controlled trial of ultrasonographic screening in pregnancy. *Lancet* 1984;2:207-11.
- Ballard-Barbash R, Klabunde C, Paci E, Broeders M, Coleman E, Fracheboud J, Bouchard F, Rennert G, Shapiro S. Breast cancer screening in 21 countries: delivery of services, notification of results and outcomes ascertainment. *Eur J Cancer Prev* 1999;8(5):417-26.
- Berwick DM, Weinstein MC. What do patients value? Willingness to pay for ultrasound in normal pregnancy. *Medical Care* 1985;23(7):881-93.
- Blanks R, Wallis M, Moss S. A comparison of cancer detection rates achieved by breast cancer screening programmes by number of readers, for one and two view mammography: results from the UK National Health Service breast screening programme. *J Med Screen* 1998;5(4):195-201.
- Blaser MJ. *Helicobacter pylori* and the pathogenesis of gastroduodenal inflammation. *J Infect Dis* 1990;161:626-33.
- Briggs AH, Sculpher MJ. Sensitivity analysis in economic evaluation: a review of published studies. *Health Econ* 1995;4(5):355-71.
- Briggs AH, Sculpher MJ, Logan RP, Aldous J, Ramsay ME, Baron JH. Cost effectiveness of screening for and eradication of *Helicobacter pylori* in management of dyspeptic patients under 45 years of age. *BMJ* 1996;312:1321-5.
- Brouwer WB, Koopmanschap MA, Rutten FF. Productivity costs in cost-effectiveness analysis: numerator or denominator: a further discussion. *Health Econ* 1997;6(5):511-4.
- Brown DW, French MT, Schweitzer ME, McGeary KA, McCoy CB, Ullmann SG. Economic evaluation of breast cancer screening: A review. *Cancer Pract* 1999 7(1):28-33.
- Brown J, Bryan S, Warren R. Mammography screening: an incremental cost effectiveness analysis of double versus single reading of mammograms. *BMJ* 1996;312(7034):809-12.
- Brown J, Buxton M. The economic perspective. *Br Med Bull* 1998;54:993-1009.

- Brown ML. Economic considerations in breast cancer screening of older women. *J Gerontol* 1992;47:51-8.
- Brown M, Fintor L. Cost-effectiveness of breast cancer screening: preliminary results of a systematic review of the literature (review). *Breast Cancer Res Treat* 1993;25(2):113-8.
- Cairns J, Shackley P. Sometimes sensitive, seldom specific: a review of the economics of screening. *Health Econ* 1993;2:43-53.
- Calam J. Clinical science of *Helicobacter pylori* infection: ulcers and NSAIDs. *Br Med Bull* 1998;54(1):55-62.
- Chamberlain J. Planning of screening programmes for evaluation and non-randomized approaches to evaluation. In: Prorok PC, Miller AB, eds. *Screening for Cancer*. Geneva:UICC, 1984:5-17.
- Ciatto S, Del Turco M, Morrone D, Catarzi S, Ambrogetti D, Cariddi A, Zappa M. Independent double reading of screening mammograms. *J Med Screen* 1995;2(2):99-101.
- Clarke PR, Fraser NM: Economic analysis of screening for breast cancer. Report for Scottish Home and Health Department, 1991.
- Cullen DJ, Collins BJ, Kristiansen KJ et al. Long-term risk of peptic ulcer disease in people with *Helicobacter pylori* infection: A community based study (abstract). *Gastroenterology* 1993;106:60.
- Danesh J, Lawrence M, Murphy M, Roberts S, Collins R. Systematic review of the epidemiological evidence on *Helicobacter pylori* infection and nonulcer or uninvestigated dyspepsia. *Arch Intern Med* 2000;160(8):1192-8.
- Day NE, Walter SD, Collette B. Statistical models of disease natural history: Their use in the evaluation of screening programmes. In: Prorok PC, Miller AB, eds. *Screening for Cancer*. Geneva:UICC, 1984:55-70.
- de Koning HJ. Breast cancer screening: cost-effective in practice? *Eur J Radiol* 2000;33(1):32-7.
- de Koning HJ, van Ineveld BM, van Oortmarssen GJ, de Haes JC, Collette HJ, Hendriks JH, van der Maas PJ: Breast cancer screening and cost-effectiveness; policy alternatives, quality of life considerations and the possible impact of uncertain factors. *Int J Cancer* 1991;49:531-7.
- deVore GR. The Routine Antenatal Diagnostic Imaging with Ultrasound Study: another perspective. *Obstet Gyn* 1994;84(4):622-6.
- Department of Health. Policy appraisal and Health. A guide from the Department of Health. London: Department of Health, 1995.
- Dew MJ. Asymptomatic peptic ulcer disease. *Br Med J* 1987;285:401.
- Dickman P, Hakulinen T, Luostarinen T, Pukkala E, Sankila R, Söderman B, Teppo L. Survival of cancer patients in Finland 1955-1994. *Acta Oncologica* 1999;38(12):25.
- Dixon MF, Sobala GM, Wyatt JJ, Axon AT. Gastric metaplasia and *H.pylori*-associated gastritis as predictors of the duodenal ulcer diathesis. *Ir J Med Sci* 1992;161(Suppl)10:11.
- Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation. A practical approach. *Med Decis Making* 1985;5:157-77.
- Drummond M, Jeffersson T. Guidelines for authors and peer reviewers of economic submissions to the BMJ. *BMJ* 1996;313:275-83.

Drummond M, Stoddart G. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford : Oxford University Press 1987.

Eddy DM. Screening for breast cancer. *Ann Intern Med* 1989;111:389-99.

Elashoff JD, Van Deventer G, Ready TJ, Ippoliti A, Samloff IM, Kurata J, Billings M, Isenberg M. Long-term follow-up of duodenal ulcer patients. *J Clin Gastroenterol* 1983;5:509-15.

Ernster VL, Barclay J, Kerlikowske K, Grady D, Henderson C. Incidence of and treatment for ductal carcinoma in situ of the breast. *JAMA* 1996;275:913-8.

Eurogast Study Group. An international association between *Helicobacter pylori* infection and gastric cancer. *Lancet* 1993;341:1359-62.

Ewigman BG, Crane JP, Frigoletto FD, LeFevre ML, Bain RP, McNellis D. Effect of prenatal ultrasound screening in perinatal outcome. *NEJM* 1993;329:821-7.

Fendrick AM, Chernew ME, Hirth RA, Bloom BS, Bandekar RR, Scheiman JM. Clinical and economic effects of population-based *Helicobacter pylori* screening to prevent gastric cancer. *Arch Int Med* 1999;159:142-8.

Finnish Cancer Registry. *Cancer Incidence in Finland 1995*. Cancer Society of Finland Publication No. 58. Helsinki 1997:9-9.

Fletcher SW, Black W, Harris R, Rimer BK, Shapiro S. Report of the international workshop on screening for breast cancer. *J Natl Cancer Inst* 1993;85:1644-56.

Forman D. *Helicobacter pylori* Infection: a novel risk factor in the aetiology of gastric cancer. *J Natl Cancer Inst* 1994;83:1702-3.

Forman D. The prevalence of *Helicobacter pylori* infection in gastric cancer. *Aliment Pharmacol Ther* 1995;9(2):71-76.

Forrest P. *Breast cancer screening: Report to the Health Ministers of England, Wales , Scotland and Northern Ireland*. Department of Health and Social Security. London: Her Majesty's Stationery Office, 1987.

Forsyth RA. The mathematics of screening mammography. *Med Hypotheses* 1988;26:21-24.

Frame PS. A critical review of adult health maintenance. Part 3:Prevention of cancer. *The Journal of family practice*. 1986;22:511-20.

Frisell J, Eklund G, Hellström L, Lidbrink E, Rutqvist LE, Somell A. Randomized study of mammography screening: preliminary report on mortality in the Stockholm trial. *Breast Cancer Res Treat* 1991;18:49-56.

Gerard K, Salkeld G, Hall J. Counting the costs of mammography screening: first year results from the Sydney study. *Med J Aust* 1990;152(12):466-71.

Gilvarry J, Buckley MJ, Beattie S, Hamilton H, O'Morain CA. Eradication of *Helicobacter pylori* affects symptoms in non-ulcer dyspepsia. *Scand J Gastroenterol* 1997;32:535-40.

Gold M, Siegel J, Russell L, Weinstein M. *Cost-effectiveness in health and medicine*. New York Oxford, Oxford University Press, 1996.

Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000;355:129-34.

- Graham DY. Can Therapy Ever Be Denied for *Helicobacter pylori* Infection? *Gastroenterology* 1997;113:S113-7.
- Greenberg PD, Cello JP. Lack of effect of treatment for *Helicobacter pylori* on symptoms of non-ulcer dyspepsia. *Arch Intern Med* 1999;159(19):2283-8.
- Hakama M. Screening. In :HollandWW, Detels R, Knox G, eds.Oxford textbook of public health, vol. 3. Oxford University Press, 1990:91-106.
- Hakama M, Pukkala E, Heikkilä M, Kallio M.Effectiveness of the public health policy for breast cancer screening in Finland: population based cohort study.*BMJ* 1997;314:864-67.
- Hakulinen T, Andersen A, Malker B, Pukkala E, Schou G, Tulinius H. Trends in cancer incidence in the nordic countries. A collaborative study of the five nordic cancer registries. *Acta Pathologica, Microbiologica et Immunologica Scandinavica* 1986;A288(94):38-41.
- Hall J, Gerard K, Salkeld G, Richardson J. A cost utility analysis of mammography screening in Australia. *Soc Sci Med* 1992;34(9):993-1004.
- Helsinki University Central Hospital. Helsinki University Hospital Price List 1999. (in Finnish)
- Hiilesmaa V, Salonen R. Äitiyshuollon seulonnan mahdollisuudet ja ongelmat. *Duodecim* 2000;47:880-86.
- Hollenberg Jim. SMLTree manual 1993.
- Holli K, Hakama M, Hakala T. Use of hospital services by breast cancer patients by stage of the disease: implications on the costs of cancer control. *Breast Cancer Res Treat* 1996;37:237-41.
- Holtzman E. Poverty and breast cancer in New York City. The New York City. Office of the Comptroller, Office of Policy Management:New York City, 1990.
- Hristova L, Hakama M. Effect of screening for cancer in the nordic countries on deaths, costs and quality of life up to the year 2017. *Acta Oncol (supplementum 9)*. Stockholm: Scandinavian University Press, 1996.
- Ihamäki T, Varis K, Siurala M. Morphological, functional and immunological state of the gastric mucosa in gastric carcinoma families. Comparison with a computer-matched family sample. *Scand J Gastroenterol* 1979;14:801-12.
- Jahkola T, Toivonen T, Tykkä H, Perhoniemi V, Tuovinen S, Roiha M.. The effect of mammography screening on the treatment of breast cancer (in Finnish).The reports of Helsinki City Department of Health. Helsinki, 1991.
- Johannesson M, Jönsson B, Karlsson G. Outcome measurement in economic evaluation. *Health Econ* 1996;5:279-96.
- Johnson S. Breast screening's bottom line, lives saved. *Administrative Radiology*, 1988;June:20-4.
- Jørgensen FS. An epidemiological study of ultrasound examinations in Denmark. *Acta Obstet Gynecol Scand* 1990;69:649-50.
- Katschinski B, Logan R, Davies J, Faulkner G, Pearson J, Langman M. Prognostic factors in upper gastrointestinal bleeding. *Digestive Diseases and Sciences* 1994;39(4):706-712.
- Kattlove H, Liberati A, Keeler E, Brook RH. Benefits and costs of screening and treatment for early breast cancer. Development of a basic benefit package. *JAMA* 1995;273(2):142-8.

Kekki M, Sipponen P, Siurala M, Laszewicz W. Peptic ulcer and chronic gastritis. their relation to age and sex, and to location of ulcer and gastritis. *Gastroenterol Clin Biol* 1990;14:217-23.

Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster V. Efficacy of screening mammography. A meta-analysis. *JAMA* 1995;273:149-54.

Kokkola A, Haapiainen R, Laxen F, Puolakkainen P, Kivilaakso E, Virtamo J, Sipponen P. Risk of gastric carcinoma in patients with mucosal dysplasia associated with atrophic gastritis: a follow up study. *J Clin Pathol* 1996;49:979-984.

Kosunen TU, Aromaa A, Knekt P, Salomaa A, Rautelin H, Lohi P, Heinonen OP. Helicobacter antibodies in 1973 and 1994 in the adult population of Vammala, Finland. *Epidemiol Infect* 1997;119:29-34.

Kuipers EJ. Review article: exploring the link between Helicobacter pylori and gastric cancer. *Aliment Pharmacol Ther* 1999 13(1);3-11.

Labenz J, Blum AL, Bayerdörffer E, Meining A, Stolte M, Börsch G. Curing Helicobacter pylori infection in patients with duodenal ulcer may provoke reflux esophagitis. *Gastroenterology* 1997;112:1442-7.

Laine L, Lewin DN, Naritoku W, Cohen H. Prospective comparison of H&E, Giemsa, and Genta stains for the diagnosis of Helicobacter pylori. *Gastrointestinal Endoscopy* 1997;45(6):463-7.

Lappalainen M, Sintonen H, Koskiniemi M, Hedman K, Hiilesmaa V, Ämmälä P, Teramo K, Koskela P. Cost-benefit Analysis of Screening for Toxoplasmosis during Pregnancy. *Scan J Infect Dis* 1995;27:265-272.

Leinikki P, Eskola J. Infektiotautien seulonta. *Duodecim* 2000;47:887-93.

Lindfors KK, Rosenquist CJ. The cost-effectiveness of mammographic screening strategies. *JAMA* 1995;274(11):881-4.

Lääketietokeskus Oy. Pharmaca Fennica 1999. (price list of Finnish drugs in Finnish)

Lääkärintarkastusohje (LTO) 1997. Puolustusvoimien Koulutuksen Kehittämiskeskus. Jyväskylä 1996.

McColl K, Murray L, El-Omar E, Dickson A, El-Nujumi A, Wirz A, Kelman A, Penny C, Knill-Jones R, Hilditch T. Symptomatic benefit from eradicating Helicobacter pylori infection in patients with nonulcer dyspepsia. *N Engl J Med* 1998;339(26):1869-74.

Meijer BC, Thijs JC, Kleibeuker JH, van Zwet AA, Berrelkamp RJ. Evaluation of eight enzyme immunoassays for detection of immunoglobulin G against Helicobacter pylori. *J Clin Microbiol* 1997;35(1):292-4.

Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. *Can Med Assoc J* 1992;147(10):1477-88.

Miwa H, Hirai S, Nagahara A, Murai T, Nishira T, Kikuchi S, Takei Y, Watanabe S, Sato N. Cure of Helicobacter pylori infection does not improve symptoms in non dyspepsia patients – a double-blind placebo-controlled study. *Aliment Pharmacol Ther* 2000;14(3):317-24.

Moskowitz M. Costs of screening for breast cancer. *Radiol Clin North Am* 1987;25(5):1031-37.

Moayyedi P, Zilles A, Clough M, Hemingbrough E, Chalmers DM, Axon AT. The effectiveness of screening and treating Helicobacter pylori in the management of dyspepsia. *Eur J Gastroenterol Hepatol* 1999;11(11):1245-50.

- Norum J. Breast cancer screening by mammography in Norway. Is it cost-effective? *Ann Oncol* 1999;10(2):197-203.
- Nyström L, Rutqvist LE, Wall S, Lundgren A, Lindqvist M, Ryden S, et al. Breast cancer screening with mammography: over-view of Swedish randomised trials: *Lancet* 1993;341:973-8.
- Okubo I, Glick H, Frumkin H, Eisenberg JM. Cost-effectiveness analysis of mass screening for breast cancer in Japan. *Cancer* 1991;67(8):2021-29.
- Paavonen J, Puolakkainen M, Paukku M, Sintonen H. Cost-benefit Analysis of First-Void Urine Chlamydia trachomatis Screening Program. *Obstet Gynecol* 1998;92:292-8.
- Page DL, Dupont WD, Rogers LW, Landenberger M. Intraductal carcinoma of the breast: follow-up after biopsy only. *Cancer* 1982;49:751-8.
- Pamilo M, Dean P, Räsänen O, Koskela J. Mammography screening in Finland - eight year results (in Finnish). *Suomen Lääkärilehti* 1995;31:3305-9.
- Parsonnet J. *Helicobacter pylori*. *Inf Dis Clin North Am* 1998;12:185-197.
- Parsonnet J, Harris RA, Hack HM, Owens DK. Modelling cost-effectiveness of *Helicobacter pylori* screening to prevent gastric cancer: a mandate for clinical trials. *Lancet* 1996;348:150-4.
- Parvinen I, Kauhava L, Helenius H, Immonen-Räihä P, Räsänen O. The hospital care and screening costs of mamma carcinoma in Helsinki, Tampere and Turku 1981-1990 (in Finnish). *Suomen Lääkärilehti* 1996;16:1804-11.
- Peeters M. Urea breath test: a diagnostic tool in the management of *Helicobacter pylori*-related gastrointestinal diseases. *Acta Gastroenterologica Belgica* 1998;61(3):332-5.
- Persson PH, Kullander S. Long-term experience of general ultrasound screening in pregnancy. *Am J Obstet Gynecol* 1983;146:942-7.
- Poses R, Cebul R, Centor R. Evaluating physicians' probabilistic judgements. *Med Decis Making* 1988;8:233-40.
- Posnett J, Jan S. Indirect cost in economic evaluation: the opportunity cost of unpaid inputs. *Health Econ* 1996;5:13-23.
- Roberts MM, Alexander FE, Anderson TJ. Edinburgh trial of screening for breast cancer: mortality at seven years. *Lancet* 1990;335:241-6.
- Roberts T, Mugford M, Piercy J. Choosing options for ultrasound screening in pregnancy and comparing cost effectiveness: a decision analysis approach. *Br J Obstet Gynecol* 1998;105(9):960-70.
- Robinson R. Effective screening in child health. *BMJ* 1998;316:1-2.
- Rosen P, Kosloff C, Lieberman P, Braun D. Lobular carcinoma in situ of the breast: detailed analysis of 99 patients with average follow-up of 24 years. *Am J Pathol* 1978;225:51.
- Saarelma O, Perheentupa J. Neuvola lasten terveyden edistäjänä. *Duodecim* 1998;47:1065-8.
- Saari-Kemppainen A, Karjalainen O, Ylöstalo P, Heinonen OP. Ultrasound screening and perinatal mortality: controlled trial of systematic one-stage screening in pregnancy. *Lancet* 1990;336:387-91.
- Saari-Kemppainen A, Karjalainen O, Ylöstalo P, Heinonen OP. Fetal anomalies in a controlled one-stage ultrasound screening trial: a report from the Helsinki Ultrasound Trial. *J Perinat Med* 1994;22:279-89.

- Saari-Kemppainen A. Use of antenatal care services in a controlled ultrasound screening trial. *Acta Obstet Gynecol Scand* 1995;74:12-14.
- Salomaa A, Kosunen TU, Aromaa A, Klaukka T, Rautelin H, Rehnberg-Laiho L, Knekt P, Heinonen OP. "Screen and treat *Helicobacter pylori* gastritis" project in a western community. *Gut* 1998;43suppl2:A112.
- Shapiro S, Venet W, Strax P, Venet L. Current results of the breast cancer screening randomized trial: the Health Insurance Plan (HIP) of Greater New York study. In: Day NE, Miller AB, eds. *Screening for breast cancer*. Bern: Hans Huber, 1988:3-15.
- Simell O, Simell T. *Neuvolakirja*. Vammala: Vammalan kirjapaino 1997.
- Simpson PR, Chamberlain J, Gravelle HS. Choice of screening tests. *Journal of Epidemiology and Community Health* 1978;32:166-70.
- Sintonen H. Milloin seulonta kannattaa taloudellisesti? *Duodecim* 2000;116:902-8.
- Sipponen P, Varis K, Fräki O, Korri U-M, Seppälä K, Siurala M. Cumulative 10-year risk of symptomatic duodenal and gastric ulcer in patients with or without chronic gastritis. A clinical follow-up study of 454 outpatients. *Scand J Gastroenterol* 1990;25:966-73.
- Sipponen P. *Helicobacter pylori*. A cohort phenomenon. *Am J Surg Pathol* 1995;19(1):30-6.
- Statistics Finland. Population by municipalities and order of magnitude 31.12.1994 (in Finnish). *Population* 1995:5.
- Statistics Finland. OSF: Population 1999;14:93-94.
- Stone MA, Barnett DB, Mayberry JF. General population screening for infection with *Helicobacter pylori*: a realistic option? *J Med Screen* 1998;5(3):162-6.
- Suomen Lääkärilehti, uutispäivävyri. Työryhmä kartoittaa geneettisen seulonnan eettisiä ja yhteiskunnallisia kysymyksiä. *Suomen Lääkärilehti* 1997;52:1576.
- Suvakovic Z, Bramble MG, Jones R, Wilson C, Idle N, Ryott J. Improving the detection rate of early gastric cancer requires more than open access gastroscopy: a five year study. *Gut* 1997;41:308-13.
- Tabar L, Fagerberg CJ, Duffy SW, Day NE. The Swedish two county trial of mammographic screening for breast cancer: recent results and calculation of benefit. *J Epidemiol Community Health* 1989;43:107-14.
- Tabar L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Gröntoft O, Ljungquist U, Lundström B, Manson JC, Eklund G. Reduction in mortality from breast cancer after mass screening with mammography: randomized trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. *Lancet* 1985;i:829-32.
- Talley NJ, Vakili N, Ballard ED, Fennerty MB. Absence of benefit of eradicating *Helicobacter pylori* in patients with non-ulcer dyspepsia. *N Engl J Med* 1999;341(15):1106-11.
- Temmermann M, Buekens P. Cost-effectiveness of routine ultrasound examination in first trimester pregnancy. *Eur J Obstet Gynecol Reprod Biol* 1991;39:3-6.
- Teppo L, Anttila A, Auvinen A, Hakama M. Syöpätautien seulonta edelleen tarpeellista ja tehokasta. *Duodecim* 2000;47:893-905.
- Thorner TM, Remein QR. Principles and procedures in the evaluation of screening for disease. *Public Health Monograph No. 67*. US Government Printing Office: Washington DC, 1961.

Torgenson DJ, Donaldson C, Reid DM. Private versus social opportunity cost of time in the demand for health care. *Health Econ* 1994;3(3):149-55.

U.S. Congress, Office of Technology Assessment: Breast cancer screening for Medicare beneficiaries: Effectiveness, costs to Medicare and medical resources required. U.S. Washington DC: U.S. Government Printing Office, 1987.

Uusimaa Health district. Uusimaa Health district health care services and prices starting from 1.1.1999. (in Finnish)

van der Ende A, van der Hulst RW, Roorda P, Tytgat GN, Dankert J. Evaluation of three commercial serological tests with different methodologies to assess *Helicobacter pylori* infection. *J Clin Microbiol* 1999;37(12):4150-2.

van Hout BA. Discounting costs and effects: a reconsideration. *Health Econ* 1998;7(7):581-94.

van der Hulst RW, Rauws EA, Koycu B, Keller JJ, Bruno MJ, Tijssen JG, Tytgat GN. Prevention of ulcer recurrence after eradication of *Helicobacter pylori*: a prospective long-term follow-up study. *Gastroenterology* 1997; 113:1082-6.

van der Maas PJ, de Koning HJ, van Ineveld BM, van Oortmassen GJ, Habbema JDF, Koos TNL. The cost-effectiveness of breast cancer screening. *Int J Cancer* 1989;43:1055-60.

van Ineveld BM, van Oortmarssen GJ, de Koning HJ, Boer R, van der Maas PJ. How cost-effective is breast cancer screening in different EC countries? *Eur J Cancer* 1993;29A(12):1663-8.

Varis K. The prevalence of ulcer disease (in Finnish). *Vox* 1988; 10:7-12.

Verbeek AL, Hendriks JH, Holland R, Mravunac M, Sturmans F, Day NE. Reduction of breast cancer mortality through mass screening with modern mammography. *Lancet* 1984;i:1222-4.

Viisainen K, toim. Seulontatutkimukset ja yhteistyö äitiyshuollossa, suositukset 1999. Stakes oppaita 34,1999.

Vintzileos AM, Ananth CV, Smulian JC, Beazoglou T, Knuppel RA. Routine second-trimester ultrasonography in the United States: a cost-benefit analysis. *Am J Obstet Gynecol* 2000;182(3):655-60.

Waitzman NJ, Romano PS. Reduced costs of congenital anomalies from fetal ultrasound: are they sufficient to justify routine screening in the United States? *Annals of the New York Academy of Sciences* 1998;847:141-53.

Wald N, Chamberlain J, Hackshaw A. Report of the European Society of Mastology Breast Cancer Screening Evaluation Committee. *The Breast* 1993;2:209-216.

Waldenström U, Axelsson O, Nilsson S, Eklund G, Fall O, Lindeberg S, Sjödin Y. Effects of routine one-stage ultrasound screening in pregnancy: a randomised controlled trial. *Lancet* 1988;2:585-8.

Warren R, Duffy S. Comparison of single reading with double reading of mammograms, and change in effectiveness with experience. *Br J Radiol* 1995;68(813):958-62.

Willke RJ, Glick HA, Polsky D, Schulman K. Estimating country-specific cost-effectiveness from multinational clinical trials. *Health Econ* 1998;7(6):481-93.

Wilson J, Junger G. Principles and Practice of Screening for Disease. Geneva: World Health Organisation, 1968

Wotherspoon AC, Ortiz-Hidalgo C, Falzon MR, Isaacson PG. *Helicobacter pylori*-associated gastritis and primary B-cell gastric lymphoma. *Lancet* 1991;338:1175-6.

Wright C, Mueller C. Screening mammography and public health policy: the need for perspective. *Lancet* 1995;346:29-32.

Yleislääkärin käsikirja 1994. Duodecim 1994, Jyväskylä.

Xia HH, Kalantar JS, Wyatt JM, Adams S, Cheung K, Eslick GD, Talley NJ. High sensitivity and specificity of a laboratory-based serological test, pulori DTect ELISA, for detection of *Helicobacter pylori* infection. *Diagn Microbiol Infect Dis* 2000;36(2):69-74.

Zavertnik JJ, McCoy CB, Robinson DS, Love N. Cost-effective management of breast cancer. *Cancer* 1992;69(suppl):1979-84.